DOCUMENT RESUME

ED 146 185 80 TM 006 542

AUTHOR Huber, Jake, Ed.; And Others

TITLE Regional Interstate Planning Project Program . . . Vol. VII. Evaluation for Effecting Educational Futures. Seminar Report.

INSTITUTION Nevada State Dept. of Education, Carson City.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Jul 76
NOTE 62p.; Report of a seminar (Denver, Colorado, July 21-23, 1976)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Administrator Education; Data Analysis; Decision Making; *Educational Programs; Educational Trends; Elementary Secondary Education; Evaluation Methods; *Evaluators; Models; Program Effectiveness; *Program Evaluation; *Regional Cooperation; School Districts; *Seminars; *State Programs; Statewide Planning
IDENTIFIERS Arizona; California; Nevada; New Mexico; Oklahoma; *Regional Interstate Planning Project; Texas; Utah; Wyoming

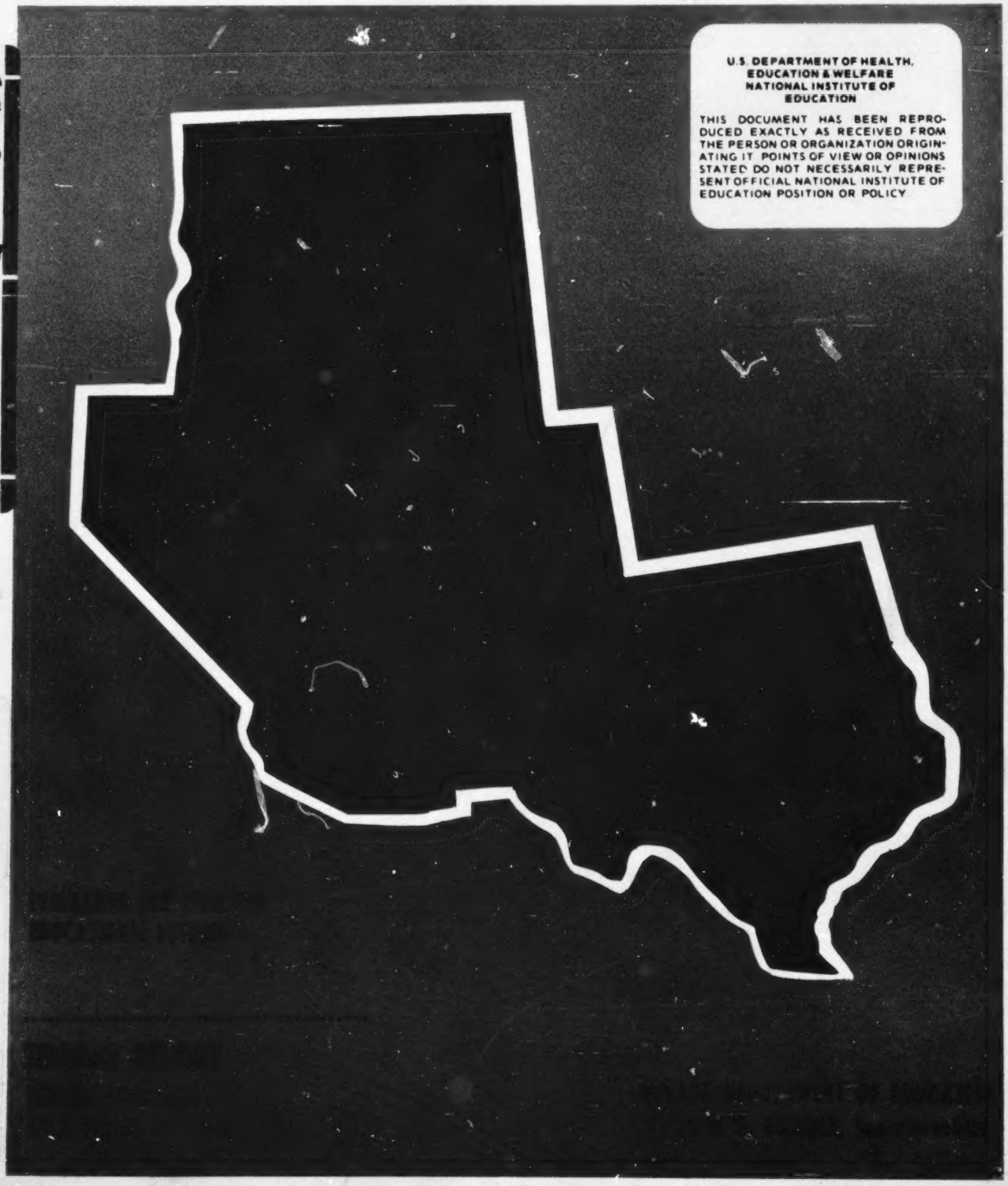
ABSTRACT

Regional Interstate Planning Project (RIPP) participants meet periodically at conferences sponsored by ten Rocky Mountain area state departments of education to discuss issues of general concern. A meeting was held July 21-23, 1976 in Denver, Colorado, to consider various means of effecting the educational future through evaluation. Excerpts from this seminar are presented. Objectives for the conference included (1) the identification of probable directions of change in education, and of ways in which evaluation could facilitate such change, (2) the sharing of experiences and ideas on evaluation activities among the RIPP states, and the dissemination of evaluation findings, (3) the review of promising or emerging evaluation practices, designs and models which may be helpful to participants in their future work, and (4) the exchange of strategies which can be used to replay, recycle and renew educational programs. Results of a questionnaire evaluating the seminar are appended. (Author/HV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

REGIONAL INTERSTATE PLANNING PROJECT PROGRAM . . .

VOL. VII



U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED146185

TM006 542

REGIONAL INTERSTATE PLANNING PROJECT

DENVER, COLORADO

July 21-23, 1976

NEVADA STATE BOARD OF EDUCATION

MRS. ROSEMARY K. CLARKE, PRESIDENT. . . . . .LAS VEGAS
MRS. SHIRLEE WEDOW, VICE PRESIDENT. . . . . SPARKS
MRS. CYNTHIA W. CUNNINGHAM, MEMBER. . . . . .LAS VEGAS
FATHER BENJAMIN J. FRANZINELLI, MEMBER. . . . . .LAS VEGAS
DR. MARY H. FULSTONE, MEMBER. . . . . .SMITH
GEORGE E. HARRIS, MEMBER. . . . . .LAS VEGAS
MRS. JOAN KENNEY, MEMBER. . . . . .LAS VEGAS
WILLIAM R. ORR, MEMBER. . . . . PIOCHE
DAVID W. HANSEN, MEMBER . . . . . RENO

JOHN GAMBLE, SUPERINTENDENT OF PUBLIC INSTRUCTION . . . . . .CARSON CITY

REGIONAL INTERSTATE PLANNING PROJECT OFFICERS

LAMAR LEFEVRE, DIRECTOR

ARIZONA	CAROLYN WARNER	BEVERLY WHEELER
COLORADO	CALVIN M. FRAZIER	ED STEINBRECHER
IDAHO	ROY TRUBY	REID BISHOP
MONTANA	DOLORES COLBURG	RALPH G. HAY
NEVADA	JOHN GAMBLE	JOHN MADSON
NEW MEXICO	LEONARD J. DELAYO	TED SANDERS
OKLAHOMA	LESLIE FISHER	CHARLES SANDMANN
TEXAS	MARLIN BROCKETTE	WALTER R. HOWARD
UTAH	WALTER D. TALBOT	WILLIAM L. HUTCHINSON
WYOMING	ROBERT G. SCHRADER	PAT WUNNICKE

The works presented herein were performed under terms of the grant from the U. S. Office of Education under the Department of Health, Education and Welfare.

An ESEA Title 5, Section 505, Project

We are an Equal Opportunity Agency

FOREWORD

The Regional Interstate Planning Project is a consortia of ten State Education Agencies that sponsors seminars on topics of mutual concern to in-service their personnel along with local education agency staff. After the seminar topic is selected, a Planning Committee, composed of knowledgeable staff from several of the SEAs, is selected to author the seminar objectives, plan the activities, select presentors, etc.

This publication contains the proceedings of a seminar on *Evaluation for Effecting Educational Futures*. To expand on the topic the following objectives were written to be achieved during the seminar:

1. Participants will be able to identify probable directions of change in education and ways in which evaluation should facilitate such change and should assist in overcoming obstacles or roadblocks to change.

2. Participants will be able to share experiences and ideas on evaluation activities among the RIPP States and to disseminate evaluation findings and reports.

3. Participants will have an opportunity to review promising and/or emerging evaluation practices, designs, and models which may be helpful to them in their future work.

4. Participants will gain ideas and concrete suggestions on utilization strategies which will increase the chances that evaluation findings will be used to replay, recycle, and renew educational programs.

In order to achieve the objectives, the following individuals were engaged to make presentations to the seminar participants, to present their evaluation models, with emphasis on the design, findings, utilization strategies and prediction regarding the outcomes: Dr. Kast Tallmadge, Vice President, RMC Research Corporation; Dr. Jim Vanecko, Senior Social Scientist and Project Director, Abt Associates, Inc.; Dr. Launor Carter, Project Director, Systems Development Corporation; Dr. Roy Forbes, Director, National Assessment of Educational Programs, Education Commission of the States.

William R. Bronson, State Director, Office of Program Evaluation Research, California Department of Education, discussed with the participants how the material for training evaluators was developed and how it was used to do the actual training.

Dr. Arthur Coladarci, Dean, School of Education, Stanford University, gave the keynote address that stimulated the participants regarding the importance of the seminar topic. He challenged us to define "Education" and the basic subjects to be taught or taken in obtaining an education. He also gave us food for thought in challenging us to be aware of where an individual will obtain the necessary information to function as a productive, worthy person in the twenty years and beyond.

Dr. Gerald Kowitz, from the University of Oklahoma, was our concluding presentor. He tied the activities, which we had conducted during the seminar, together and gave us advice on how to plan in the future to improve our project activities.

In addition to the above, I would like to thank Pat Watson, Oklahoma, Bea Smith, Texas, and Jim Miller, New Mexico, for their forthright statements during their reaction regarding the evaluation models. To the representatives from each of the States who made the presentations during the Pack and Brag sessions, to the State Coordinators for their assistance in planning, coordinating, moderating, taking pictures and helping in many other ways, your help and assistance is appreciated. To Verl Snyder and Harry Phillips, thanks for the informative, helpful briefing regarding education legislation. To all, a big thank you for a job well done.

Lamar LeFevre
Director
Regional Interstate Planning Project
Nevada State Department of Education
4055 South Spencer, Suite 234
Las Vegas, Nevada 89109
(702) 385-0191

INTRODUCTION

The Regional Interstate Planning Project participants meet periodically at conferences sponsored by ten State departments of education to discuss new or topical issues of general concern. The most recent meeting was held July 21-23, 1976, at Denver, Colorado, to consider various means of effecting educational future through evaluation.

The R.I.P.P. Conferences are financed with funds provided under the Elementary and Secondary Education Act of 1965, Public Law 89-10, Title V, Section 505 as amended. The project is administered by the Nevada State Department of Education under the supervision of Superintendent John Gamble who serves as Chairman of the Policy Board. Other cooperating State departments of education are Arizona, Colorado, Idaho, Montana, New Mexico, Oklahoma, Texas, Utah, and Wyoming.

This particular issue was plagued with problems resulting from recording equipment failure during the conference. In most cases the speakers were gracious enough to provide us with assistance in recapturing their presentations. However, a few were not able to provide them in time for publication. The editors apologize for this and will make every effort to avoid similar occurrences in the future.

During the process of transcribing, editing, and compiling the R.I.P.P. proceedings, the editors have sought to keep the material both brief and informative. The resulting report is, hopefully, a useful record of many thoughtful discussions.

Dr. Jake Huber, Co-Editor
Dr. Evalyn Dearmin, Co-Editor
Dr. Len Trout, Co-Editor
Research and Educational Planning Center
College of Education
University of Nevada
Reno, Nevada

TABLE OF CONTENTS

	PAGE
RIPP CONFERENCE PARTICIPANTS	*vi*
PHOTOGRAPHS	*vii*
SUMMARY--EVALUATION: SANCROSANCT AND PROFANE *(Dr. Arthur Coladarci)*	1
ASSESSING THE COGNITIVE ACHIEVEMENT BENEFITS RESULTING FROM ESEA TITLE I PROJECTS: THE PLAN FOR 1977 AND BEYOND *(Dr. Kast Tallmadge)*	2
RESEARCH ON DEMONSTRATION COMPENSATORY EDUCATION: STUDYING IMPLEMENTATION AS A GUIDE TO POLICY *(Dr. Jim Vanecko)*	5
CALIFORNIA EVALUATION AND IMPROVEMENT PROJECT *(William Bronson)*	8
SUSTAINING EFFECTS OF COMPENSATORY EDUCATION ON COGNITIVE GROWTH--A NATIONAL STUDY *(Dr. Launor Carter)*	11
NEWS FROM THE POTOMAC *(Harry Phillips)*	16
A DISCREPANCY -BASED PLANNING MODEL USED IN NEVADA *(James Kiley)*	18
ASSESSMENT MODEL FOR VOCATIONAL EDUCATION IN NEVADA *(John Madson)*	20
STATEWIDE EVALUATION OF INSTRUCTIONAL PROGRAMS IN NEW MEXICO *(Alan Morgan)*	22
EVALUATION EFFORTS IN THE STATE OF TEXAS--CAREER EDUCATION, DEVELOPING READING AND MATHEMATICS OBJECTIVES, EVALUATING COMPENSATORY EDUCATION *(Dr. Walter Howard and Carl Defibaugh)*	25
UTAH'S STATE ASSESSMENT PROGRAM *(David Nelson)*	27
COMPREHENSIVE-SYSTEMATIC PLANNING IN WYOMING *(Thomas O. Kraner)*	29
EVALUATION OF CAREER EDUCATION IN ARIZONA *(Richard Ruff)*	31
INVOLVING ELEMENTARY AND SECONDARY TEACHERS IN DATA UTILIZATION *(Bill Connett)*	33
EVALUATION ACTIVITIES OF THE PLANNING, RESEARCH, AND EVALUATION SECTION, OKLAHOMA DEPARTMENT OF EDUCATION *(Charles Sandmann)*	34

PAGE

INFORMATION UTILIZATION IN DECISION-MAKING
(Dr. Roy Forbes) 36

SUMMARY, EVALUATION AND IMPLEMENTATION
(Gerald Kowitz). 40

EVALUATION. 44

REGIONAL INTERSTATE PLANNING PROJECT
CONFERENCE PARTICIPANTS

Denver, Colorado

July 21-23, 1976

ARIZONA

Jerry George
Don Kearns
Richard Ruff
Beverly Wheeler
David Williams

COLORADO

Bill Brown
Dwight Cool
Cathy Felknor
Roslyn Grady
Ed Steinbrecher
Bob Willard

IDAHO

Reid Bishop
John Briggs
Donald Carpenter
Bert Stoneberg
Helen Werner

MONTANA

Bill Connett
Ralph G. Hay
Don Moore
Phil Ward

NEVADA

James Bean
Richard Gunkel
James Kiley
LaMar LeFevre
John Madson
Courtney Riley
Shirlee Wedow

NEW MEXICO

Allen Ferrel
Joyce Goodwin
James Miller
Alan Morgon
Roberta Stuart

OKLAHOMA

Gladys Dronberger
Milton Ford
Frank Hobbs
Victor Kehn
Gerald Kowitz
Bobby Martin
Charles Sandmann
Pat Watson

TEXAS

Carl Defibaugh
Walter Howard
Ken Kyle
Bill Marshall
Bea Smith

UTAH

Richard Burbidge
Bill Cowan
William Hutchinson
Nola Miller
David Nelson
Don Richards

WYOMING

Mark Fox
Tom Kraner
Marcia Troxel
Pat Wunnicke

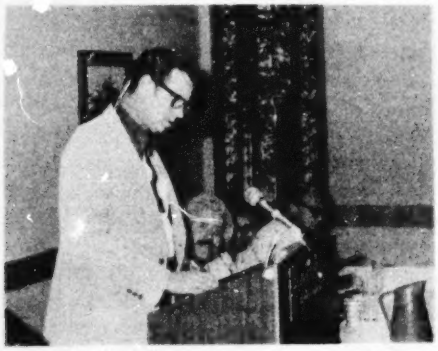
GUESTS

Bill Bronson
Launor Carter
Arthur Coladarci
Betty Feilbert
Joe Lawter
Verl Snyder
Kast Tallmadge
Len Trout
Jim Vanecko
Roy Forbes
James Mortensen

Verl Snyder, USOE Representative for the RIPP Project, discusses Interstate Cooperation.

Dr. Kast Tallmadge, RMC Research Corporation (back right) makes a point with interested participants about achievement benefits resulting from the Title I Projects.

Jim Kiley, Nevada, makes a point regarding data utilization in educational matters.



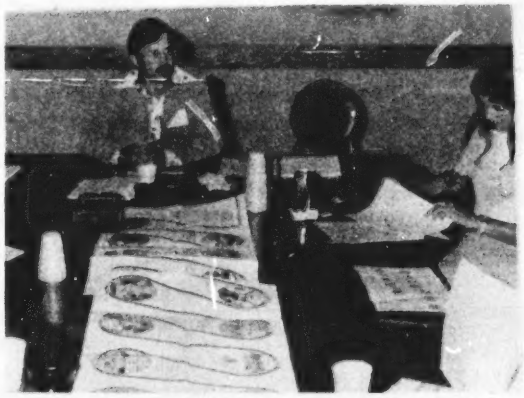
Dr. Mark Fox tells the participants how the Wyoming State Education Agency uses evaluation data to inform Legislators.

Dr. Launor Carter, Systems Development Corporation, emphasizes a point regarding effects of Compensatory Education on Cognitive Growth.

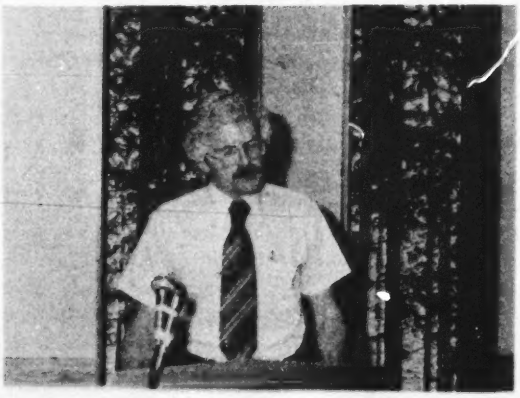
Verl Snyder and Dr. Harry Phillips from the USOE keep participants abreast of Education Legislation being considered in Washington, DC.

Dr. Beverly Wheeler, Arizona, out in front of the Arizona Delegation.

Montana Delegation - Phil Ward (Center) Montana Coordinator.



Montana shows Evaluation Models.



Dr. Roy Forbes, National Assessments of Educational Programs, Education Commissions of the States, discusses how to utilize data to make decisions.

Wyoming Delegation - Pat Wunnicke, Coordinator (third from left).

Colorado presents its study and operational procedure for concept six-year around school programs.

LaMar LeFevre, RIPP Project Director, opens a session.

Dr. Ed Steinbrecher, Arrangements Chairman and RIPP Coordinator, stands on the right of the Colorado Delegation.

Dr. Jim Vanecko, Abt Associates, chats with Pat Watson after presentation.

Big John Madson (back row center) gets the Nevada Delegation together.

Utah Delegation - William Hutchinson (second from right) State Coordinator of RIPP.

Don Richards, Utah, shares an Evaluation Model with other State participants.

Idaho Delegation
Dr. Reed Bishop (second from left), Idaho Coordinator.

Four Major Presentors (left to right): Gerald Kowitz, Verl Snyder, Launor Carter, and William R. Bronson, State Director, Office of Program Evaluation and Research, California Dept. of Education.

Dr. Walter Howard (standing), Superintendent Jim Miller, Pat Watson, and Bea Smith react to the Conference Major Presentation.

Dr. Gerald Kowitz, Conference Evaluator (second from left), Gladys Dornberger, Member of the Planning Committee (third from left), and Pat Watson of the Oklahoma City Delegation.

Oklahoma shows Evaluation Models with other States.

Dr. Walter Howard, RIPP Coordinator from Texas, in the middle of the Texas Delegation.

SUMMARY

EVALUATION: SACROSANCT AND PROFANE

by

Dr. Arthur Coladarci
Dean, School of Education
Stanford University

The logics, tactics, and techniques of "evaluation" are not invariant across situations. Rather they vary with the kinds of purposes held in view. At least these differing sets of purposes arise in education:

(a) Evaluation to make decisions about the adequacy of administrative, organizational structure and process.

(b) Evaluation to make decisions about pupils (e.g., placement, programming).

(c) Evaluation to make decisions about adequacy of the instructional programs (curriculum and teaching method).

We often confuse the three, thinking of all as constituting the same evaluative task.

Focusing on the third purpose, another distinction should be made: (a) evaluation for the purpose of "testing" the program (summative evaluation), (b) evaluation for the purpose of improving the program (formative evaluation). In the first (summative), the canons of "science" apply--that is we hold "sacred" the principles of reliability, objectivity, reproducibility, etc. However, it is counterproductive to extend this "scientism" rigidly to the second (formative) mode. Rather, in continuous evaluation for continuing program improvement, we should learn to tolerate the "less scientific" as long as the information is the best that can be obtained under the program conditions.

ASSESSING THE COGNITIVE ACHIEVEMENT BENEFITS RESULTING FROM ESEA TITLE I PROJECTS: THE PLAN FOR 1977 AND BEYOND

by

Dr. Kast Tallmadge, Vice President
RMC Research Corporation, Mountain View, California

In all of our assessment models we use posttest scores as a measure of post-treatment performance, and you're free to select a test which you feel measures what it is that you're trying to teach. We try to generate some scores on the total test which we think they would have gotten if they had not participated in the project. In all three models we use this same concept--all use posttest scores as a measure of post-treatment performance. They differ in terms of how they generate the no-treatment expectations.

The three models are labeled A, B, and C. Model A is a norm-reference model. It probably looks like what most people are doing today. Model B is a control group model, and in most Title I projects it is probably illegal. Model C is what we call a special regression model. We've developed a slight modification of that which we call a regression projection model. We use that to generate a no-treatment expectation over a particular set of circumstances.

In Model A we use published, normative data. We use it in the same way that we use control-group data. We treat the students as if they were a special case in the control group. Instead of comparing kids against "the national norm" we are comparing treatment kids against kids in the normative sample who got the same pretest scores. We generate this low treatment expectation in a very simple way. We assess at pretest time the percentile standing of kids with respect to the national norm. We then assume, all other things being equal, they will maintain their status with respect to their grade-level peers without special help over the pre- and posttest time period. In other words, if they start out in the 20th percentile and you don't help them, they will end up at the 20th percentile. Many people jump at this expectation saying, "Then why do our kids fall farther and farther behind the national norm?" This is true. They fall farther and farther behind the national norm because they stay at the same percentile. They fall farther behind in raw score points, they fall farther behind in standard score points, they fall farther behind in grade-equivalent scores, but they maintain their same percentile. Nationwide they have to. It can't be any other way. What we are saying is that they stay somewhere between the kids that are just a little bit brighter than they are and kids who are just a little bit dumber than they are. We assume that unless we can help them, they will maintain their pretest percentile status at post-test time. So, then, we can get a no-treatment expectation from the expected posttest score. That's our no-treatment expectation.

If you want to use this norm-reference model, you have to do it with standard scores and percentiles and not with grade equivalents. You have to use percentiles that were empirically derived, in other words, percentiles which were not extrapolated by the test publishers. The California Achievement Test, for example, was normed in February. So the percentile norms of the California Achievement Test in February are empirical norms. The test publishers present

in their manuals all the spring norms in addition. But those norms are constructed by linear interpolation between Februaries, under the assumption that project growth is linear over the nine months of the school year, with one-third as much gain being made over the three summer months. That assumption gives you a nice linear growth over the artificial ten academic-month year, but it is simply untenable.

It's particularly untenable with disadvantaged and gifted kids because disadvantaged kids decline in their growth rate over the summer and gifted kids exceed their growth rate. In fact, if you were to take data from any test publishers' norms and plot the growth, i.e., get a student at a standard deviation below the mean (the 16th percentile) and one at a standard deviation above the mean (the 84th percentile) and an average student at the 50th percentile and plot them over the school year and then over the summer and then over the next school year, you will see that all three grow very much the same over the school year. The difference between them occurs during the summer. The disadvantaged kids fall farther behind, the national norm kids do what they are supposed to do, and the gifted kids do a whole lot better than they do while they are in school.

I'm not assuming that the different kids in the norms are in different programs. I'm just saying that you have to use empirical norms that set some additional requirements as to when you can test. If you want to use the California Achievement Test, it only has empirical norms in February. That is the only time you can test your kids and make valid comparisons against the norms. You can't test in the fall or spring because the norms are projected and the errors in projection are sufficiently large to invalidate the norm reference of comparison.

We were able to convince the California Achievement Test publishers that their projected norms are not adequate. They are systematically in error. The disadvantaged kids were systematically too low in the fall and systematically too high in the spring. The California Achievement Test is currently being renormed. We take some modest credit for that. The CTBS falls in the same category. For educators who want to use a fall pretesting and spring posttesting model at the present time, only two standardized achievement tests can be used in conjunction with the norm-reference model.

One additional requirement that we make in conjunction with Model A, our norm-reference model, is that you cannot use a single set of test scores to select the kids who are going to be in the treatment and to pretest. If you use the same measures to select kids because they have low scores and then use those scores for pretest measures, you have random error working in your favor. You can get statistical regression toward the mean which adds up to whatever treatment effects you might have had, and you get too big a mean. We require that you select the kids first by whatever method you want. Once you have selected them, then you must pretest them.

Model C is quite complex. The positions of each student are plotted on two axes, one axis for pretest scores and one for posttest scores. Whenever we pretest and posttest a group of kids and then plot their positions in this manner, we will get an array that looks like a scatter plot. It is so characterized because it is elliptical in shape and the cases tend to cluster on the seven. If we look at any particular test score and make a slice that corresponds to a pretest score, and we look at the midpoint of that slice, we should

have exactly half of the cases from that score above the line and half below. If we look at the distribution of scores, the distribution of posttest scores, and the kids that got this pretest score, we should have something that approximates normal distribution. The mean of that distribution is the midpoint of the slice. Any place we make a slice, we should secure the same results. Or, if we draw a line between the midpoints of all the slices, we secure a regression line. Basically the regression line is the line that goes through the midpoint line of all the slices, and in all cases it gives us the best prediction of the posttest score at a pretest score position. The symmetry at this array is what makes the special regression model work.

With this model we can predict posttest scores for hypothetical kids by knowing pretest scores. If they have a certain pretest score, then they should get a predictable posttest score. Suppose we administer the pretest to 1,000 kids, but we only have enough money to serve 300 of them in a special program. We want to obey the law and serve the neediest ones; so we're going to try to get the lowest pretest score until we get 300 kids. There we will draw the line. Only kids below that line get the Title I treatment. None of the kids above get the treatment.

Models B and C both require testing non-participants. This is the main objection to them. Another objection to Model B is that you can't really find kids comparable to Title I kids unless they are in a single semester-long program. Model C is also fairly complex in terms of calculations, and some people have objected to it for that reason. Some people object to Model A because of the restrictions on testing time--double testing in order to select any pretest kids. And there is some stigma attached to using national norms.

The Educational Amendments of 1974 state that the Commissioner shall provide such technical and other assistance as may be necessary to the State educational agencies to enable them to assist local educational agencies in the application for and the development and systematic evaluation of programs in accordance with the models developed by the commission. This assistance is going to take two forms during this coming year. First there will be a series of 11 three-day workshops, one in Washington, D. C., for the "feds," and one in each of the regions, primarily for State-level evaluator people, State Title I people, and State Title I evaluators. Each of these workshops is intended to train people to become conversant with the system and also to prepare them to conduct similar workshops for the LEA's within their States.

The second form of technical assistance will be technical assistance centers--again, one for each of the regions. These centers are to provide essentially free consultant services to the States and the local agencies within each region. The States will call upon these agencies to conduct workshops, to help them with workshops, and to help with planning and proposals to those centers.

RESEARCH ON DEMONSTRATION COMPENSATORY EDUCATION: STUDYING IMPLEMENTATION AS A GUIDE TO POLICY

by

Dr. Jim Vanecko
Senior Social Scientist and
Project Director
Abt Associates, Inc.

The research I am here to describe to you is a study in distributional equity. It is a study which departs from traditional educational research but which could not be more in the center of the research tradition. It is a study derived from the hypothesis that if compensatory educational services are distributed on the basis of educational disadvantage (achievement test scores), rather than on the basis of poverty, then more children from poor families will receive compensatory services and poor children will receive more services. Perhaps the hypothesis is facetious, perhaps merely provactive, but it is interesting and testable.

If one considers that current regulations play havoc in desegregated school districts so that money is allocated to schools because of the poor kids who live nearby but go to school elsewhere, if one considers that a poor child who can't read but lives on the east side of the street may be denied services because many children who are also poor but can read live on the west side of the street, then it is not an unreasonable hypothesis. We social scientists call it the ecological fallacy.

Educational equity has been much discussed in recent years. Traditional research on the issue can be characterized as following two approaches, the accounting approach and the Coleman approach. Both suffer from exactly the same problem--they don't deal with teaching, with process, with implementation, with service delivery, with what actually gets to kids. They don't deal with the organizational constraints or the delivery of equal educational opportunity. The accounting approach defines equity in terms of inputs--the distribution of goods and services. The Coleman approach (named after James Coleman) defines equity in terms of outputs--the achievement test outcomes of education. Both in their worst forms are black box approaches. The fact is that the concern of educators and educational researchers alike is with the educational process. The process requires inputs of resources, and it must be judged in terms of outcomes; but to focus on either or to assume one simply produces the other is a mistake. We have taken care to avoid that mistake. Only time will tell if we have succeeded.

Background and Objectives

Title I has received much criticism, particularly in the areas of program administration, student selection, and service delivery. In an attempt to address these, Congress authorized the National Institute of Education (NIE) to conduct a major study of compensatory education. NIE drafted a plan for a comprehensive set of research projects. The Study of Demonstration Title I Compensatory Education Projects, one of this set of research projects, deals

with the effects of using educational need rather than economic need as the basis for distributing Title I services. Sixteen Local Education Agencies contracted with NIE during the 1975-76 school year to devise plans for implementing Title I programs based on an educational definition of need. Those plans accepted by NIE will be implemented during the 1976-77 and 1977-78 school years.

The research portion of the Demonstration Projects is being conducted by Abt Associates Inc. (AAI). This research will compare the distribution of instructional and other educational services during the 1975-1976 school year, under the current federal regulations, with the distribution during the 1976-1977 and 1977-1978 school years, with federal regulations waived. It will provide the basic analysis for possible modifications in the Elementary and Secondary Education Act (ESEA) of 1965 with regard to allocation policy: eligibility criteria policy and distributive policy.

The definition of the study within the set of studies planned by NIE, and the districts' Demonstration plans have led to the establishment of five major research objectives. They are presented below in order of priority.

1. To estimate and analyze the distribution of *services received* by students. To achieve this objective a log is completed by teachers for a sample of students for entire sample days. This log will provide data on the full range of instructional services received by different types of students: Title I versus non-Title I students, students from varying socio-economic and ethnic backgrounds, and students of varying achievement levels.

2. To estimate and analyze the distribution of *services delivered* across types of schools and types of students. To achieve this objective questionnaires are completed by principals and teachers. These questionnaires focus on changes in programmatic content and strategies, as well as distributional changes.

3. To estimate and analyze *community reaction* to the Title I programs and the Demonstration. To achieve this objective interviews are conducted with parents of Title I children and members of Parent Advisory Councils.

4. To estimate and analyze the *costs* or savings associated with the altered allocation and eligibility policies.

5. To gather descriptive information on the possible achievement outcomes of the Demonstrations.

Given the research questions enumerated above and the preliminary district plans for the Demonstration, the sample was designed to include four types of schools:

1. Schools which received Title I during 1975-1976 and which were expected to receive Title I during the implementation years of the Demonstration (1976-1977 and 1977-1978).

2. Schools which did not receive Title I during 1975-1976 but which were expected to receive it during the implementation years of the Demonstration.

3. Schools which did not receive Title I 1975-1976 and which were not expected to receive it during the course of the Demonstration.

4. Schools which received Title I during 1975-1976 but which were not expected to receive it during the implementation years of the Demonstration.

Once schools have been selected, all third and fourth grade classrooms within these schools are included in the sample. Two students (if possible, one Title I and one non-Title I student) are randomly selected from these classrooms along with two alternates for each primary student. Alternates are involved in the research only if the primary student moves, is sick during the data collection period, or is unavailable for some other reason.

Because the research calls for information on the regular and, if appropriate, Title I instruction in reading and mathematics delivered to these students, teachers who provide this instruction are included in the sample. Because the research calls for information on the organization of the districts' schools, the services they provide, and the students they serve, all principals in the districts are asked to participate in the study. Finally, since the research requires information about parent involvement in and reaction to the Demonstration, selected parents and Title I Parent Advisory Council members are asked to participate in the research.

Basic Analytic Model

The basic analytic model underlying the Title Demonstration Study consists of groups of factors outlined in the objectives earlier and the links between them that will be analyzed in this research.

The process leading from eligibility and distribution policy decisions to program changes and, ultimately, to changes in the students served as segments of the model.

We hope that at the end of this study, we can say what gets to kids as a result of different allocation policies. The policies are the choice of the districts participating. What gets to kids is for us to determine by focusing on implementation and the process of service delivery. If these hopes are realized, we will be ready to say something about what is reasonable policy and, more importantly, what policy alternatives affect kids.

CALIFORNIA EVALUATION AND IMPROVEMENT PROJECT

by

William Bronson, State Director
Office of Program Evaluation Research
California Department of Education

One of my basic assumptions is that evaluation is too important to leave to the local evaluator. We have to get data from the village watchman because that is where we usually get it--the problem is that the village watchman just puts down what he damned well pleases. What we need to do is train the watchman.

That is one of the features we have in our project--because the charge from ESEA Title IV, Part B, was to raise the skill at the local level in the area of program planning and evaluation dissemination. Until program evaluation practices are accepted and implemented from the classroom level to the school board level, there will be little change in programs made as a result of evaluation.

If any of you read the Rand report on how decisions are made, I think you will be very discouraged about going into evaluation for decision-making purposes. However, I think people need to learn how to make decisions based upon evaluation of data, and perhaps one of the reasons that they have not been using it as adequately as they should have is that it has not been presented in a form that seems to them to be useful. Evaluation must be seen as something beneficial to the program and not as an additional task done for someone else. Evaluation is gathering information and doing assessments for the purpose of making decisions--for the purpose of doing something with the information. Assessment is merely looking, measuring, and seeing exactly what something is. We have a State assessment program, but it is not an evaluation program. We are trying to tell the local districts it isn't enough. What you need to do is develop for yourselves an evaluation program which will give you the information you need in order to make decisions, in order to improve your program. We must remove evaluation from the realm of mystery and intrigue, take it out of higher statistics, and present it as a useful tool for policy-makers and decision-makers.

I'd like to differentiate here between evaluation information and experimental research design. Evaluation information is secured from whatever or whoever can be helpful to me as I make a decision--the classroom teacher, the principal, etc. I'm not looking for ultimate proof. I am not looking for an experimental research design which is aimed at ultimate truth. The level of generalization makes a difference. If I know this information is going to be helpful to me, that's all I need. If I'm going to make some other, wider generalizations; then I have to be much more sure of what I am talking about in terms of reliability, validity, and all the rest. Then the purpose is quite different. We need to identify very clearly the purpose for which we are doing an evaluation. If we don't, then we are going to be criticized on the basis of experimental research design when perhaps that is not what we were intending. I was very happy

to hear Dr. Coladarci comment that unreliable information was not immoral. Unreliable tests are acceptable when they offer information that is useful. Just because a test does not come up to the goals or reliability standards, it still may be very helpful in finding out specific things. The use needs to be identified and accepted for what it is and not extrapolated into something it is not. I direct my remarks here really at program managers and teachers.

Long-range interests are going to be best served when professional evaluators are able to secure appropriate, timely, accurate information that the policy and decision-makers can understand and trust. But many evaluation reports are automatically suspect when they come out of State departments. Until we can communicate in such a manner that can be understood and believed, then we might as well be doing something else. And in order to secure this timely and accurate information, we need to go back to the village watchman level, to the local level.

Gathering information at the local level is very difficult. Perhaps the professional evaluator plays a role as an interpreter between the classroom teacher and the decision-maker. In a sense, he needs to be bilingual. He needs to talk the language of the teacher--the level at which the information is collected--and he needs to talk the language of the decision-maker. The person who is going to write the reports for the decision-maker from the information gathered in the classroom is going to have to understand each of these languages sufficiently well to pick up mere ounces of difference.

The federal evaluator is very important in this role, and we are trying to enhance his role as we conduct our workshops. As you well know, it is not going to be possible to develop successful evaluators in a short time. Much training must go into it. We offer a three-day workshop, but if you can turn out a statistician or an evaluator in three days, I would like to know about it. We have not been able to do that so far in California. What we think we are doing is taking program managers, planners, and administrators who have skills in these areas, and we're trying to sensitize them to the problems of evaluation so that when they see them in their process of planning curriculum, they recognize them as problems of evaluators. They are not scared away from them, and they at least know how to handle them or when to ask for help. That really is the essence of our training program.

The structure of our workshop is such that we are dealing primarily with small districts. The larger districts are just beginning to come to the workshops. When ESEA, Title IV, was first inaugurated, its objective was to raise skills at the local level in the area of planning, evaluation, and dissemination. We began with the assumption that the money would dry up eventually. When it dried up, we wanted to leave a residue of expertise around the State that would still be available to help.

We went to our county offices of education first. They have been in business for some time and will probably be in existence long after the funding of our project. If we could train these people to continue our work, they would then be able to work with the district and could train another cadre of people to work at the level of the local districts. We identified eight county offices and designated two of these as development centers--the other six became satellite centers. The primary role of the development centers was to develop

the actual workshop itself. The primary emphasis of the satellite center was on the dissemination, the actual conducting of workshops. The State's role was coordination.

Let me tell you just a little bit about the development centers and the people who were employed as directors. We said to the county superintendents, "Get us people who know their constituents, who can talk to the people and are respected. We really don't care about their evaluation skills." As a result we got people in there who were oriented toward administration, towards curriculum, toward planning--mostly curriculum and planning--and this was the group that developed the material. They decided what they needed; then they went to the professional evaluators and said, "Hey, how does this sound? Does this framework contain anything that could possibly be subsumed under the total area of evaluation?" Consequently, each of these centers is a separate unit. Some were developed in the north and some in the south. Although our format is not completely identical, we have set up a coordinated planning system. Each unit is an entity in itself; however, each needs to be a part of the whole.

You can enter our total plan at any point and move in any direction to cover the entire context. If a person comes in to you and says, "Hey, I want a test for this," you don't say, "I'm sorry, we can't give you a test until you go back there because that's where our project starts." We have to be able to respond to the need and say, "Testing is only part of the picture. Here is the rest of the picture. Here is what you're asking for." Then we move in those directions and show them there is more to evaluation than testing, more to evaluation than writing reports, more than statistical analysis, more than needs assessment. It is the sum total of all of these things. Until you get all of these elements, you are not going to have an evaluation program. Here is where you are. Now I am going to help you with this.

In our workshops we started out with three basic steps: plan, conduct, and use. Well, that is simple enough. I think that all people can see that you have to plan it, you have to conduct it, and you should be able to use it. What does the plan consist of? Well, you have to determine the evaluation purpose and requirements: Why are you doing this? If you understand why you are doing it, it is much more clear what you have to do. If you are evaluating merely to fulfill federal or State requirements, it's very simple what you have to do. If you are evaluating for public relations purposes, that restricts what you need to do. If you are evaluating to improve programs, you have a different focus.

These kinds of determinations are very important as you look at our program planning. You can't really separate program planning and program evaluation. When you look at it in this light, should we use this kind of a format for secondary education reforms? The planning requirements are there, and they are there quite specifically. Once you find out what you are doing it for, then you determine an evaluation plan and procedures. You have to obtain your assessment tools, you have to log the data, you have to analyze the data, you have to report results, and then hopefully you should apply the findings because if there is no application of the findings, then you shouldn't have gotten that far in the first place.

SUSTAINING EFFECTS OF COMPENSATORY EDUCATION ON COGNITIVE GROWTH--A NATIONAL STUDY

by

Dr. Launor Carter, Project Director
System Development Corporation, Santa Monica, California

Congress asked the Office of Education some time ago to find out the answers to a number of questions plaguing them with regard to Title I. Congress said, in effect, "Hey, look. We've been appropriating a billion and one-half to two billion dollars a year for Title I, and the evaluation studies so far tend to be somewhat disconcerting. There just isn't a great deal of evidence that the Title I programs have been terribly effective in picking up the disadvantaged children." So Congress asked O.E. to do a defending study to tell them how effective Title I really is. Secondly, Congress said, "We want to know whether Title I funds are being properly distributed--whether they should be based upon economic disadvantagement or whether they should be based upon educational disadvantagement. Investigate that problem for us also." So these are two of the major things that System Development Corporation is attempting to do. Between ourselves and the Office of Education we have tried to formulate a number of policy questions which we hope will be answered by this particular study.

These are some of the different policy questions we are trying to answer. First are some descriptive questions. For the country as a whole, what is the nature of compensatory education? What are the things that are being done in classrooms that are called "compensatory"? Second, who receives what kind of compensatory education? We know what the regulations say about who should receive it, but who, in fact, receives it? What kind of children are they? What are their educational needs? What kind of economic disadvantagement do they have? What kind of racial or ethnic composition do they have? Questions of that nature.

Third, how can one improve the effectiveness of compensatory education? We are going to be studying many different kinds of compensatory education, and hopefully, at the end of the study we will be able to say these kinds of programs seem to be more effective than other kinds of programs. Fourth, what effect does the receipt of compensatory education have on educational growth? This is a question of--if the child starts out at grade one at a disadvantage, what is the growth from grade one to grade two, to grade three, or to grade four, and so on? Is there a quick catching up that is held, or does it fall off? How does that relate to the different kinds of students that are receiving the programs?

Fifth, what is the optimum duration of compensatory programs? Some people argue that you ought to have your compensatory programs quite early in first or second grade; others say compensation ought to occur later. Where should you compensate if you have a limited amount of money? Sixth, what happens when student programs are discontinued? If a student happens to be in a compensatory program under Title I and then advances to a certain level of achievement, he may be ineligible for the program. What happens when he no longer gets the program? Does he fall back down again, or not?

Seventh, at what grade level do compensatory programs seem to be most effective? Eighth, how do home and local circumstances affect the child? Do the intellectual, school-oriented characteristics of the home, seem to have an effect on the child and how? Finally, what seems to be the influence of summer school on the achievement growth of the child. For some children summer school or the lack of summer school serves to be a depressant; for others it seems to pick them up. So, those are the policy questions that hopefully, at the end of this study, we'll be able to say something quite definitive about. As a result of what we are able to say, we believe that Congress will change the Title I laws. From a broad educational point of view, then, this could be a very important study.

Now, what is the overall design of this study? We shall address these nine policy issues that I have just described and we shall define the kind of population in schools we are going to investigate. The schools will consist of three or four different kinds. Our first study will be longitudinal; that is, we shall follow over several years the students in 251 schools, but not the schools and not the classrooms. These 251 schools have been selected in such a way that we are quite confident they are a truly representative sample of all the schools in the United States. We are going to follow all of the students in the school, not just the Title I students, not State program students, but all the students, whether they are disadvantaged or not, because we want to get a comparison between the disadvantaged students and the other students. We are taking grades one through six, but we are going to try to follow the students as they leave the sixth grade and go into seventh, eighth, or whatever, to see what happens to them as they leave the elementary grades and perhaps the Title I program.

In addition, there is a group of 30 schools where--as far as the records tell us at least--there is a high level of disadvantagement. These schools are in poverty districts, but they have no compensatory programs or very low compensatory programs. You might think under Title I regulations such a circumstance would be impossible, but it appears that, indeed, there are such schools. So we have identified 30 of them and these will also be a part of the longitudinal study.

The sampling plan involves a selection of these schools in terms of geography (they are distributed throughout the country) and in terms of the size of the LEA from which the school comes. We really wanted to get at school size, but it turns out there is no national record of school size. There is no place you can go and say, "Hey, what is the size of Emerson School?" You can go and say, "What is the size of this LEA?" So we had to get LEA size to substitute for school size, and then develop a poverty index. We now have all 64,000 schools in the country categorized by the level of poverty of their LEA.

The instruments we are using for students have to do with measures of academic achievement, measures of functional literacy, measures of attitude and self-concepts. We are also collecting data on the school program. We are trying to find out about the instructional programs offered to these students, and the characteristics of the school personnel involved in rendering these educational services--the teachers, principals, etc.

Since we are going to follow each student we must identify him or her year after year without ever knowing the student's name. After all, there is a privacy law--the Buckley Amendment--which prevents us from using the student's

name. All we know is the student's number. And that means that the records and the match between the name and the number have to stay at the school. Conversely the school will never have access to the individual data provided by the teachers or the principal or others because we retain that data. In order to follow the students individually, we're making up packets for each one so that the teacher can fill out the test and instructional information about every student. Everything is contained together in that one packet. At SDC we have to mail these at the right times to meet the school's schedule, which may begin anywhere from as early as August 13 to as late as September 20. These are some of the things we have done during our initial planning year.

The second year, the year we are in right now, is our first operational year. Beginning this fall, in the third week of the school year, we will administer the student test materials. They will be shipped back to Santa Monica where we will score the materials very rapidly in our own scoring facilities, and we will return to each school the achievement test results for each student so that the schools can use them for their own counseling or evaluation purposes. During the school year we will be collecting information about the teacher, the principal, and the program. In the spring we will re-administer the materials, and hopefully, start analyzing our data to see what kind of academic rope we have associated with each of the different programs and instructional feature characteristics.

We will do the same things in the second and the third operational year, and if Congress and God are willing, in the fifth and sixth operational year--all the time analyzing our data of course. In the seventh year we will come up with the answers to our initial policy questions. We will try to answer these questions as we go along within the overall design of the longitudinal study.

The most important parts of this study depend upon determining a truly national representative sample. How do you get that? We sent out a questionnaire to a sub-sample of 5,054 principals of those 64,000 schools which met our definition of the kind of schools we wanted. We asked 13 different questions concerning the size of the school, the number of teachers, the poverty level in that school, the ethnic composition in the school, the kinds of compensatory programs in the school. We wanted that information so we could check against our base sample of 54,000, but also so we could check our longitudinal sample of 251 schools against the bigger sample of 5,000. Over 95 percent of the questionnaires have been returned by now.

Now the question is, how do you select the schools from among this large sample? Well, we took these three criteria--geography, LEA size, and poverty index--and made up a table with cells. We wanted a school from each of the ten O.E. regions. We had three sizes of LEA's--small, medium, and large, believing that this distribution would be sure to secure us a fair percentage of rural schools, medium schools, urban schools, and some large city schools. To establish the poverty index, we over-collected on high poverty level LEA's, and then selected from some medium-poverty levels; then, from the rest of the LEA's we set up 90 cells, and took all the schools in the country and put them in those 90 cells, and then drew randomly from the schools that fell into those cells. That ought to have given us 270 schools, but you will notice we really had 251 because some of our cells turned out to be zero. For example, in Region XIII there is no large LEA that has a high level of poverty. In some places there are no rural schools with high levels of poverty. The characteristics of the schools were fed into the computer which generated the numbers. The schools

selected have been generally very cooperative. Of the 251 schools that we identified, only 16 declined to participate, and of course we just took the next pool on the cell.

Now, what about instruments? On the achievement tests we are giving a part of the reading section and a part of the math section of the Comprehensive Test for Basic Skills. For our purposes, that was the best instrument of all the instruments published. Actually three or four of them were almost the same, but we could get from the CTBS test publishers information on the ethnic composition of the norming groups, and we could do a deep biasing study on the individual items for that test better than we could for some of the others.

As you know, there has been much criticism of achievement tests. To measure achievement, other than by the formal achievement test, we had a panel advise us on functional literacy tests appropriate for second, third, and fourth graders. It turns out that there are no such tests; so we have devised a functional literacy test which the students, at least, think is great. It has good psychometric characteristics. Unfortunately it correlates too highly with the achievement test. After all, if you are going to be measuring reading as the kids actually use it out in their everyday play activities, or their work activities, or arithmetic, it is going to be generally the same sort of arithmetic and reading that they have in the achievement test. But it is within the context of the playground, of the soda parlor, of the kitchen--how you make Bisquick pancakes and things like that.

We also have a survey of attitudes toward math, of attitudes toward reading, of attitudes toward the school and toward the self as a student. It is partially a standard survey published by a psychological corporation and partially an instrument which we developed ourselves. Those are the instruments we are going to be giving to the students, and we think they are pretty good instruments. The pretest showed that they worked well.

The next one, for student participation and activities, is the weakest part of our study. That particular instrument is supposed to figure out for each student how much reading and what kind of reading instruction or math instruction that student was exposed to throughout the year. Four times throughout the year we are asking the teacher to fill out this information for each student. In a typical week, how many hours of reading, how many hours of math instruction do the students have and how was it given? Was it the teacher dealing with the whole class, was it a teacher aide working with that particular student, was it pull-out instruction, was it a high intensity language lab--what was it? Hopefully, we will find out student by student what kind of instruction in math and in reading they did get. In the pretest that instrument worked very well in some of our schools, but in some of the schools it was not a satisfactory instrument because it seems that in some of the schools in the country, the teachers in the elementary grades have great difficulty reading. We've revised the instrument extensively, and we're trying it out again. As we get the results, we may devise another instrument because that is a very important part of the study.

We're asking the teacher to fill out something about the student's background, the nature of the home they come from, whether they went to a different school before they entered this school, sex, etc. We are trying to find out from each student what they did last summer. We want to know if they went to summer school. Did they go to a camp, did they just stay at home, did they work in

the field, what did they do during the summer? Then we're completing a compensatory education log for each student. Are they supported by Title I or some other title or State education funds or special district funds?

We're also collecting information about the nature of the school staff, the size of the school, the location of the school, the amount of funds going to the school, things of that nature. The principal questionnaire has demographic information on the principal, but more importantly the principal's attitude toward discipline, toward relations with his teachers, toward freedom to be flexible in their teaching, etc. For the teacher, in general, we are asking the same questions: demographic information, educational philosophy, procedures--what kinds of activities are preferred, how math and reading are handled.

For program characteristics we are trying to identify what each of the compensatory programs strive for, what they consist of. We want to know if they have any other means of assessing what actually transpired during the period of compensatory instruction? Is there any defined objective?

When we started our study, we wanted to be sure that we had in our sample some schools that offered what were thought to be the best compensatory programs in the country. We sought a contrast. We asked State departments of education, the Office of Education, and various consultants to nominate schools that they thought had exceptionally good programs. From the 250 schools nominated, a panel we selected weeded that down to about 80 schools which we then visited to try to verify what kind of program they had. We now have 45 schools that appear to have different, successful compensatory programs. These will be treated just as any other part of the sample from the point of view of data collection, but they were not randomly selected--they were purposely selected. Some of them are not Title I schools. Some of them are in very affluent school districts. This selection confuses people who think ours is a Title I study. Well, yes, it is; but it also has a lot of other little handles on it. So, now we have in our sample 326 schools and about 120,000 students.

Another part of the study is cost-effectiveness. We have sub-contracted with RNC to do a cost effectiveness study on the compensatory programs. We will be collecting at the classroom level information about the resources that are used in that classroom and what the student gets from our participation law. We are going to try to relate the amount of resources used for instruction to the progress made by the students.

We have also sub-contracted with DECIMA, a survey research organization in Santa Ana, to seek an answer to that question which Congress asked--what is the relationship between education disadvantagement and economic disadvantagement? We want to know what the correlation between these two is and whether to advise a change from economic to educational disadvantagement in the allocation of funds. What influence will that have on the distribution of funds to the States? To a politician, this is a very important question. DECIMA will be going to 15,000 households of the students involved in the study, the homes of approximately 50 students from each school, trying to find out three different kinds of information. What is the income and economic level in that home? What is the intellectual climate of the home? What is the parents' attitude toward school?

As you can see this is a big and complex study, and because such surveys are often criticized for getting superficial data, we will conduct a verification study by using semi-structured techniques for two weeks at each of these schools. Hopefully, we will then be able to describe those programs and communicate them to the education community.

"NEWS FROM THE POTOMAC"

by

Harry Phillips, Legislative Liaison
U. S. Office of Education

The educational measures undergoing deliberation during this session of Congress deal primarily with higher education and vocational education. All appropriations for FY '77 are under discussion right now. The likelihood of a veto is strong because just the educational section alone of the HEW appropriation is about three-quarters of a billion dollars above the Presidential request. Generally Congress wishes to dispense with the measures now before it in order to take a thorough look at future alternative roles for the federal government. In my judgment the next Congress will be seriously deliberating upon a new role in education. Changes will be made in the priority of support for education, and Congress will consider major alternatives to the current method of allocating federal resources. The categorical grants-in-aid approach has burgeoned into more than the educational community can cope with. Consequently, there will be some fundamental shifts in the national policy on education.

We still have many problems to work out in relation to the fiscal year beginning date shift from July 1 to October 1. All of the dates for the submission of program reports and fiscal reports have now been changed. In some instances State educational agencies may need to keep dual sets of fiscal accounts and make dual program accounting arrangements until the transition smooths out a year from now. The change becomes critical in a few instances in which report dates alter ingredients of formulas. In the vocational educational legislation, in the Title I legislation, and indirectly in the Emergency School Aid Act legislation, annual per pupil expenditures of the State and local resources become a factor in interpreting how much money flows out to local school systems or to the State. If you change a fiscal year from July 1-June 30 to October 1-September 30, you have some basic terminology to redefine or some basic accounting adjustments to make in terms of a record-keeping system that generates annual per pupil expenditures. We hope that this data can continue to be collected and maintained on July 1-June 30 basis rather than changing the statistics to conform to the new fiscal year. Under the FY shift, recording of program data may become more complex, but retaining the old fiscal year collection period is preferable to changing the definition of basic data items.

Other items of current interest are the educational authorizing bills that are now before the Congress. In vocational legislation there are considerable variations between the House bill and the Senate bill. Both bills enhance the future of vocational education, but the way the two Houses envision that future differs considerably. The greatest difference is in the way vocational education should be governed at the State level. The House projects a continuation of the current governance pattern. In some States the decision-making body is the State Board of Education, and in some States, it is the State Board for Vocational Education. But, the House has enlarged the role of that State board by asking that its membership be broadened and that it contact a wider-ranging group of people for consultation.

The Senate version would upset the existing State arrangement considerably by introducing a new planning commission for vocational education--a commission

that would approve the way a State decides to allocate and utilize its vocational education resources that are Federally appropriated. The commission would not be the State Advisory Council on vocational education either; it would be an entirely new State Planning Commission. This substantial change is highly controversial. How it will be resolved is unpredictable at the moment; however, most of the discussants I hear oppose the Senate version in favor of the House version.

The House version of the vocational education bill contains considerable consolidation; the Senate version is a continuation of the existing vocational legislation. The Senate bill also contains an extension of the higher education authorities currently in operation, but the House bill does not contain that. Some of these items are a bit controversial. In addition, the President has conceived an anti-busing piece of legislation that may be introduced and be superimposed on the Senate version as it comes to the floor. With the hurried schedule Congress is now in during an election year, plus the fact that the Senate is caught up right at the moment on an extensive tax reform bill, we may get very fast decision-making and compromising on this legislation.

The next few months are going to be very busy for this Congress. The educational decisions they make will be vitally important. There are a few pieces of fringe-area legislation that State educational agency personnel should be alert to because there may be a few benefits for education. One is a bill that the Senate passed over the President's veto. This bill has to do with the public works jobs. The potential educational benefits are that the bill will allow the refurbishing, the rebuilding, or the revamping of public institutions that were, in the first place, built in part with federal resources; and that provision applies to an array of educational buildings.

The CETA legislation--the Comprehensive Employment Training Act--is for reauthorization, and will be enacted. The new CETA bill may have expanded authority in it. Another bill, the general revenue sharing bill, will be renewed sometime between now and December 31. The former version of that bill prohibited using more than half of the funding that goes to local levels for purposes of education. That prohibition will probably be removed from the new bill.

A DISCREPANCY-BASED PLANNING MODEL USED IN NEVADA

by

James Kiley

Several years ago in the State of Nevada the Governor commissioned a study on educational accountability. People in communities all around the State had an opportunity to present their views on education--their likes and dislikes, aspirations and hopes, etc. The final report consolidated the attitudes of teachers, administrators, parents, interested citizens, and students. The legislature then took that report and asked the Department of Education to work with the school districts in implementing some of the recommendations and to report back to the legislature for the succeeding two years the progress that had been made in the State.

Given that directive, we developed a comprehensive planning system model which we call a discrepancy-based model. The model takes into consideration three different elements in the needs assessment: student needs; student performance, program or process variables in education; and societal expectations in education. First we establish process objectives or program objectives, student performance objectives, and educational goals as perceived or translated by the society. Then we attempt to measure the present status against those statements of expectation. By comparing the difference between expectation and present status, we identify educational needs.

The complete model is conceptualized as a wheel because planning isn't sequential or linear; it seems to be circular with all the parts interrelated. Evaluation is related to every segment of that wheel. In going through the planning process, one is continuously involved in feedback, adjustment, and evaluation. The model is based upon what we call process objectives and performance objectives. The performance objectives indicate changes of behavior on the part of learners; and teachers or administrators can be learners as well as students.

The strategy that we used to incorporate this comprehensive planning system into all the federal programs administered by the department was to ask districts to complete a needs assessment based upon the model, and to follow the steps and the comprehensive planning process as they developed a proposal and reported against it. We developed application and reporting forms that were consistent with the model. If they were interested in the planning model, we asked them to allow us to field test the model in one, or two, or three of the schools, and to give us one full day with their teachers to help get the project started. We informed each district superintendent of the ground rules. We emphasized that nothing was being forced on them, if they didn't want to do it, they didn't have to, and no one would feel bad about it. We asked them to talk to their building principals, and if they were interested in having their staff participate in the plan, fine; but if they were not, we would go right on down the road.

We did ask the district to make a commitment. If they wanted to participate, they had to free their administrative staff and their teachers for a full day so that we could meet with them during school time to describe the process to

them. Although we would provide the technical assistance to them throughout the field test of the planning system, the district could back out at any time.

The teachers got together and established performance objectives for youngsters. We gave them four different sets of materials to look at that were really program structures or hierarchies of skills and concepts which had been developed by others, but we didn't pressure them to adopt any particular system. We let them identify the hierarchies of skills and concepts they felt were best representative of their instructional program. After a month or two we would return to discuss and critique what they had put together, working toward the development of an instructional management system in reading and arithmetic that really became for them a new pupil-monitoring system and a new pupil-rating system.

When we first met with the teachers they were confused and upset about having to become involved in the project, but when they started talking to one another and found out that some of them really did not know how to teach reading, that they did not know the discrete skills and concepts that some of their colleagues pursued in the teaching of reading; then they began to get interested in the entire process. We encouraged discussions among teachers at the same grade level and at different grade levels. We asked them to cut across grades. We tried to practice what we were preaching. When we met with the teachers, we would set specific objectives to be met by the next time we returned.

As we tried to get teachers to look at skills and concepts and how youngsters demonstrate these skills and concepts, one interesting outcome was that teachers began to get a better appreciation of evaluation. Not only did they want to know whether Johnny or Mary had mastered a skill, but also whether or not their colleagues were doing a good job in measuring those same skills. Now they are getting interested in putting together their own homemade CRT.

The instruments are not completely sophisticated, of course, but the teachers are able to talk about what they want the youngster to do. They are able to think about what kind of setting they can create to let the youngster demonstrate that skill. And we are bringing in consultants from larger school districts or from private corporations to expose the teachers to more sophisticated strategies and procedures that they can adopt as part of their instructional management system. Presently we are not moving as fast as they would like us to.

The Department of Education is conducting statewide evaluations through contracts that we award to outside agencies to perform different tasks for us. Last year, we awarded a needs assessment contract for a right-to-read project to three outside agencies. We were able to involve three agencies because the model for needs assessment is clear enough so that we can communicate among contracting agencies and don't have to depend upon one agency to complete the entire needs assessment. We asked the contractors to measure student performance in reading at the third, fifth, seventh, twelfth grade levels and at young adult levels in the State of Nevada to attempt to define for us functional literacy and give us some standard that we might be able to use in discussing functional literacy with our constituents in the State. Then we wanted to look at process variables and requested one contracting firm to visit the exemplary reading programs in the State of Nevada and identify for us the various instructional components of those exemplary reading programs. Finally we tried to find out how well prepared our teachers are--what kind of training and experience do the teachers who work in these exemplary reading programs have. We have tried to evaluate both process and student performance this past year through a discrepancy-based model which has proved effective.

ASSESSMENT MODEL FOR VOCATIONAL EDUCATION IN NEVADA

by

John Madson

In Nevada, we have a Statewide planning system for setting up vocational education programs. The system incorporates five major categories: needs and opportunity information, planning, student services, instruction, and educational accountability. Within each one of these categories are 12 functions. Every application or every local plan--whether the district is starting a new program or is renewing on an annual basis an existing program--must address each function.

In order to provide the best vocational education program, planners must first determine what the community needs are. Without a needs assessment as a foundation, the planning can be haphazard. In addition, providing the community an opportunity to participate in an assessment brings about positive community reaction and support. The needs assessment acts as a barometer. If students are not getting jobs because they are poorly prepared or because they are trained for jobs which are not available, then the whole purpose of vocational education is questioned.

To avoid the many pitfalls in planning and administering various programs and to keep abreast of community needs, I believe that the assessment model which we put together for use by the local education agencies provides the kind of information needed for planning and implementing the various programs. A good needs assessment model should define the systematic process which describes what is and what ought to be. The process entails population analysis, job market analysis, job performance analysis, and identification of resources available. From an assessment of these areas, the priorities of the program are developed by each agency.

Because teachers must implement programs, their experience, ability, interest, and needs must also be taken into consideration. In Nevada, once a year teachers are asked to complete a questionnaire which gives input into the program planning. This information also provides a barometer for inservice training.

An assessment of the general population requires information about how many people live in the area to be served, what is the age breakdown, what is the female-male breakdown, what percentage represents minorities, disadvantaged, handicapped, what are the housing characteristics, what is the level of education, what percentage of the population is bilingual, etc. Much of this information is available in various census studies, but the idea is to summarize the information and make it readily available.

Because the job market is constantly changing, the local administrator has to keep his finger on the pulse of the employment situation; and this can best be accomplished through a job market analysis. Specifically, the job market analysis is designed to answer the question, "How many and what kinds of jobs should students be training for?" In spite of the relative mobility of people, a great many students stay in their local community, and work and live there; so the job market analysis of the district is extremely important.

Then we consider the area, whether it is rural or metropolitan. The type of area is especially significant in planning community college or continuing programs. We also take a look at the State and national employment needs. State and national priority items, such as solar energy and air purification, are important factors in the planning of vocational education programs. Then we consider the job performance analysis. The job performance analysis is probably the most important element in the entire planning model. If the students are not trained in the proper competencies, they will be unable to get the jobs for which they are being trained. The job analysis addresses these questions: Are the tasks required by the employers identifiable components in the vocational education program? Have existing students or ex-students obtained the sequence of skills necessary to enter the job, and to what degree have they achieved these skills? Is there an appropriate committee of teachers, administrators, and advisory members--citizens and students--to review and update those skills needed in the occupation? Is this procedure carried out on a routine, regular basis?

A job performance analysis also covers existing programs to see if they are meeting the needs of people within the district. By reviewing annually the national definition of the particular occupation or job, and collecting data from employers regarding the required competencies to perform the job, the analyst can find out which competencies should be incorporated into the program.

Nevada is moving rapidly toward competency-based education, particularly in vocational education. We will no longer graduate students or give them passing grades if they are not competent in the basic skills of the occupation for which they are being trained. We have a form to obtain information concerning competencies from various employers. We interview target groups of students to see that the skills they have learned compare with the skills that the employer requires. Finally, we do a follow-up study, tracing the student after he graduates.

The final, major category in the model is resource analysis. In this part we identify the physical resources available. We also look at the federal, State, and local finances that are available, and we make an effort to look at the human resources available in communities--volunteer teachers and aides and speakers. We encourage interest through the use of local lay people on advisory councils. We also consider the availability of the professional staff.

This, very briefly, constitutes the needs assessment model that we provide for all the local school districts in Nevada. When they start a new program, they go into much more depth than when they continue a program; but each year, the administrators, the local vocational director, and the principal of the school look at the vocational education program and assess what the program is accomplishing and what it should be accomplishing.

STATEWIDE EVALUATION OF INSTRUCTIONAL PROGRAMS IN NEW MEXICO

by

Alan Morgan

The New Mexico Statewide evaluations consist of several measurements including a State-required standardized testing program and a New Mexico objective-reference test program. Five years ago we went to one-third of the school districts in the State of New Mexico and asked teachers, administrators, and parents of students to help us write some basic cognitive-skill objectives. From these we derived a set of objectives that appeared to be the most frequently chosen objectives in the State. The end product of that survey was a basic Cognitive-skill Objective Bank, which was divided into the subject areas of mathematics, science, social studies, and communication skills.

Then we returned to the school districts in the State and held objective-rating workshops. In a forced concensus process, the participants were asked to make a distinction between those skill objectives which were of most concern and those which were of least concern. After the school district had chosen a set of instructional objectives, it became the responsibility of the State education agency to develop tests for those various schools, based upon the objectives they had selected, in order to allow a rural school district to choose different kinds of instructional objectives from those of a large urban school district.

Over the last several years we have been looking at what kinds of objectives districts have chosen, and we have compiled and published these in a document entitled Trends and Priority Objectives as Selected by New Mexico School Districts. Interestingly enough, although we have found some variation in the kinds of objectives that are being selected, generally there is some uniformity in the concerns that people have all across the State of New Mexico.

In addition, we analyzed objective selections in terms of role types. For example, did teachers choose different kinds of objectives than did students, or did students choose different kinds of objectives than did administrators, etc. We found no statistically significant difference in the kinds of objectives that were being selected by each of those role types. It would appear that there is a great deal of commonality in the kinds of concern people have throughout the New Mexico school districts.

Whenever a district chose objectives, the State Department of Education chose a test to measure those objectives and reported that information back to the school. The purpose of doing so was to help reassure those who believed tests developed in New York or California could in no way measure the kinds of progress or skills New Mexico children seem to have or do not have. What we were doing is sampling more items in a certain domain--the division of two of three-digit numbers--in the objective-based test than we could do in a standardized test; and furthermore the standardized tests normally don't cover the same areas, or at least the same precise areas, as we could cover in the objective based test. You won't find too much about the various cultures in New Mexico in the CTBS, nor will you find out too much about career education. The CTBS is not sensitive to local change. The whole process of developing a test, field testing with a subset of the population, revising that instrument,

printing, packaging, and distributing is incredibly difficult and complex. If we could have done it any other way, we would not have gotten into the test construction business because it is not something that amateurs do very well. We've termed these objective-based tests rather than criterion-reference tests for the simple reason that we did not require a student to meet an objective by completing X percentage of items correctly.

We also have dealt with regression analysis for the last four years, and we report back to the schools how they did as compared to some sort of a range of expectancy as computed by the State department of education. We use three factors to compute that expectancy as a pre-primary input factor to project some sort of a range of performance of each school in the State. We use previous test performance, median family income, and student mobility. If the school consistently scored below that expectancy, then the State department of education would become more involved with that school. So we are using the system not only as a way to identify trends and priority objectives but also as a way to deploy staff of the department of education.

There are other facets of the Statewide evaluation program, including the reporting of college entrance test scores. Although these are not part of the State-required testing program, we try to provide some information about why the scores change as they do.

When we first started the standardized testing program, teachers often didn't get a chance to see the teacher's manuals that came with the tests--or if they did, the manuals were handed out at the time the test was given and taken up immediately afterward. So we developed a New Mexico teacher's manual for the New Mexico classroom teacher. The manual deals with questions like, "Why are we doing this? What do we get back? How can we use this information?" Any teacher who gives a Statewide test gets the teacher's manual to keep.

Teachers are probably not too excited about having to give a test from which they do not get back much information directly at the classroom level. Now that the evaluation program is becoming more a responsibility of the local school district, we assume that there can be a much better application for the classroom teacher and maybe for the individual child than in the past.

Five years ago, educators in New Mexico at almost any level--classroom teachers, teachers' aides, administrators, superintendents of various school districts--were not really thinking very "objectively." Most of our program planning was somewhat primitive in terms of the Minimum Standard for Public Schools in the State of New Mexico, a State board document with which all schools have to comply. The Minimum Standards essentially boil down to three statements: the schools are required to assess the needs of their students, to plan a program to meet those needs, and to evaluate that program. Minimum standards imply some regulation of schools and districts to insure "quality" education. What is quality education? What are those basic skills that every kid should have? We're groping more with that problem than we have in the past. With the movement of the objective-based testing program to the local level, we are now considering a minimal-skills assessment program, although the State Board is undecided on the question of minimal skills or minimal competency testing.

In New Mexico fifty or more percent of the students are children who have an ethnic background different from the majority white community. In testing programs there are no special exceptions for the bilingual child, exclusive of

those that are set out in the regulations; that is, if a student cannot deal with the English language, if that's not his dominant language, if he doesn't understand the testing situation that he or she has been placed in, then the student is not required to complete the test. General guidelines indicate that students who are in the mainstream programs are to take the tests at the grade level as they are required. If the student cannot deal with the testing situation because of a linguistic variation, then we do not require the student to participate. Usually that doesn't happen very often.

Unfortunately, even though we have had bilingual programs in New Mexico for a considerable length of time, we still do not have an abundance of non-English materials for a number of reasons. There are variations of dialects across the State, and some of the Indian languages are not in written form. Nine percent of the population is Indian, which constitutes a fairly numerous group of students. We have in recent years conducted only an informal evaluation of bilingual education programs in the State.

The needs assessment process that we went through for the entire educational system began at the State level and moved to the district level. When we started developing objectives and asking districts to make a decision about which objectives were more important, we were getting into the assessment process. We've come to understand in the last couple of years that the process needs to start at the grass roots level instead.

Because the new minimum standards, passed in July of this past year, require schools to assess student needs at the school level, we put together a process guide for their use. They could use other kinds of models for needs assessment, but we simply wanted to provide some examples of some of the kinds of instruments that might be helpful. In essence this document offers a way a school district can identify student needs, and that's really our basic philosophy. From that determination will evolve other kinds of needs--staff needs, program needs, facility needs, perhaps even policy needs.

Two years ago a State mandate required that by 1977-78 every school district in the State would offer a State-funded early childhood education program. It was not a change of the compulsary school attendance law; we didn't say every kid of a certain age had to be in the program. We simply said that every school had to offer a program different from kindergarten. Our intent was not to have a mashed-down first grade. A rather rigid, classical instruction design would not be appropriate for the five-year-old child. We wanted to emphasize social and emotional growth more than intellectual or cognitive growth. We have also prepared for schools a document dealing with the evaluation of early childhood education programs.

EVALUATION EFFORTS IN THE STATE OF TEXAS-- CAREER EDUCATION, DEVELOPING READING AND MATHEMATICS OBJECTIVES, EVALUATING COMPENSATORY EDUCATION

by

Dr. Walter Howard
Carl Defibaugh

Editor's Note: Dr. Howard discussed briefly Texas involvement in the development of Essential Reading and Mathematics Objectives and the Texas effort in the assessment of career education. Both of these efforts have produced publications. Essential Reading and Mathematics Objectives is currently available from the Texas Education Agency, and the project report, Career Education Measurement Series (CEMS), will be published shortly.

Mr. Defibaugh discussed the Annual Report of Programs for the Disadvantaged in Texas (Title I Regular, Migrant, and State Compensatory Education). He also reviewed briefly the following evaluation materials prepared by various school districts and the Texas Education Agency:

From Austin Independent School District:

1974-1975 Final Report: Individually Guided Education (IGE) Program
1974-75 Final Evaluation Report, ESEA Title I Project
1974-1975 Final Report: Cook Elementary School
1974-75 Final Report: ESAA II Pilot Project Assist
Final Evaluation Report: ESAA Bilingual/Bicultural Project, 1974-75
Final Evaluation Report: ESEA Title VII Bilingual Project, 1974-75
School Based Planning Manual CIPO

From Dallas Independent School District:

Measurement Profiles, 1975-1966
A Summary of the Major Findings from "Everybody Knows What Profit Is"
Abstracts of Research and Evaluation Reports, 1974-75
A Plan for the Use of Research, Evaluation, and Information Systems Resources, 1975-76

From Fort Worth Independent School District:

Short Term Assessment of the Middle School Plan
Initial Assessment of the Intensified Learning Plan
Short-Term Desegregation Effects: The Academic Achievement of Bused Students, 1971-72
1974-75: The Evaluation of Six Kindergarten Curricula
Abstracts, 1974-75: Department of Research and Evaluation
A Comparison of "Continuous Progress" and "Traditional" Classes in Three Middle Schools
Final Report of the Title IV, Part A: American Indian Needs Assessment and Referral Program, 1974-1975
Assessment of the Utilization of the Time-Out Structures

From Houston Independent School District:

The 80-Minute Period
ESAA Basic Program: Final Report, 1975-76
Contemporary Learning Center Final Report, 1975-76
Community High School Final Report, 1975-76
An Educational Program Audit for QIE-MS Programs Final Report, July 1976
Addendum to Final Report: Auditors' Subjective Reports, QIE-MS, July 1976

From Texas Education Agency:

Self-Evaluation Guide for Local Districts for Vocational Education of Handicapped Students
The Development of Evaluative Criteria for Training the Handicapped: A Summary of Project Activities
Guide to Proposal Evaluation for Vocational Education of the Handicapped
Texas Education Product Study: Digest of Final Report
A Model for Evaluating Programs in Vocational Education for the Handicapped
Evaluation of Coordinated Vocational-Academic Education in Texas
Evaluation of Career Education Projects in Texas: Houston, Harlandale, and Fort Worth Independent School Districts
Special Project Career Education: Texas Education Agency

UTAH'S STATE ASSESSMENT PROGRAM

by

David Nelson

The 1975 Utah Statewide Educational Assessment was developed and refined as a joint effort of the Utah State Board of Education and the Institute for Behavioral Research in Creativity. The study was under the general direction of Dr. Don K. Richards, Administrative Assistant and Planning. Dr. David E. Nelson, Specialist in Evaluation, had specific responsibility for the project within the State School Office. Dr. Robert L. Ellison directed the project for the Institute for Behavioral Research in Creativity.

The 1975 Utah Statewide Educational Assessment was Utah's first comprehensive educational assessment. This statewide program of educational assessment was conducted because Utah educators are very concerned with improving educational practice. Statewide assessment is one example of the commitment of teachers and administrators throughout the State, local school districts, and the Utah State Board of Education and its staff to provide a quality education for each student within Utah's public school system.

The study was based on the goals approved by the State Course of Study Committee in the Fall of 1972 for public education in Utah. Aspects of five of these goals were measured in the 1975 Utah Statewide Educational Assessment: Intellectual Maturity, Emotional Maturity, Social Maturity, Aesthetic Maturity, and Productive Maturity. The students' Enjoyment of School was also assessed as it was implicit in all of the Utah Educational Goals. Besides measuring student outcome in these areas, the study measured many factors which might influence student achievement, including demographic characteristics, such as socio-economic background, ethnic group membership, and sex; as well as school classifications, such as school size and location. Another extremely important part of the study was the measurement of key aspects of the educational process, including measures of techniques and strategies used by educators to help students learn, for example, Classroom Participation, Development of Career Talents, and Expectations.

The 1975 Statewide Assessment effort involved the testing of random samples of both elementary and secondary schools. Over 4,000 5th grade students in 67 elementary schools and over 3,000 11th grade students in 26 high schools took part in the study.

In the Intellectual Maturity goal area, Utah students were superior in mathematics and science achievement, when compared to a national norm group. The Utah students were clearly outstanding in science achievement at both the 5th and 11th grade levels, but the public schools should be encouraged to find ways of helping all students reach higher levels of achievement in mathematics and science. This is especially crucial for students from lower socio-economic backgrounds.

In the Emotional Maturity goal area, Utah students had generally high levels of positive adjustment on each of the scales in this study. The chief area of concern was the lower level of Academic Self-Concept reported by the 11th grade students when compared to 5th grade students. This suggests that specific

instructional strategies aimed at enhancing levels of Academic Self-Concept of all students, with particular emphasis on secondary students, should be included in the curriculum.

In the Social Maturity goal area, fewer than 20 percent of the Utah students sampled at each grade level indicated they felt they were either "a little weak" or "rather poor" on a variety of peer relations activities. Thus, the great majority of students at both grade levels indicated their competence in dealing with peer relations was at least "very good." The results also indicated that more opportunities for Utah students to engage in leadership experience are needed. Also needed are more opportunities to engage in teamwork activities and other group activities.

In the Aesthetic Maturity goal area, 11th grade students generally performed at higher levels than the 5th grade students. However, performance was not high at either grade level. The State School Office Art and Music Specialists, who assisted in the construction of these scales, felt that ideally the content measured by these tests should be mastered by students close to the termination of the students' elementary school years. Much could be done to enhance student achievement in both the visual arts and music. Such efforts might consist of several different specific strategies, including emphasis of pre-service and in-service training, innovative programs, greater use of curriculum guides, etc.

In the Productive Maturity goal area, the standardized Career Planning Knowledge test showed that Utah 11th grade students scored about the same as the national average. Another important measure of a student's Productive Maturity was a scale called Independent Development. This scale measured the extent to which students participated in self-initiated learning activities which increased their general competence. In general, on selected activities, Utah 5th graders reported higher levels of independent learning than did 11th grade students. The present thrust in the State in the area of career education seems warranted in view of these results. Certainly, Utah students could benefit from increasing familiarity with the nature of the world of work, as well as additional emphasis on actual job experience. Many aspects of career education might well be integrated into almost every segment of the existing curriculum.

For Enjoyment of School, reasonably positive results were obtained at both grade levels, with the 11th grade students scoring slightly higher than the 5th grade students. However, the possibility existed that this result was due to the attrition of students who disliked school. The significant relationships of this scale to educational process measures suggest that broader implementation of specific instructional strategies would lead to higher levels of Enjoyment of School.

The results obtained from the educational process measures included in this study--Individualization of Instruction, Reinforcement of Self-Concept, Classroom Participation, Development of Career Talents, and Expectations--indicate that such processes warrant increased levels of application in Utah schools.

An important finding of the study, and one which has been found nationally, was the impact of socio-economic status on the achievement and attitudes of Utah students. This finding suggests that steps need to be taken to deal with the educational disadvantages which accrue to students from lower socio-economic backgrounds.

COMPREHENSIVE-SYSTEMATIC PLANNING IN WYOMING

by

Thomas O. Kraner

Educational evaluation in Wyoming is directed toward the definition of school goals, assessment of the school environment, assessment of the children's response to the environment, feedback of information for relevant decision-makers, decisions for modification of the school's environment and then a recycling of the whole process. Philosophically, much of it is focused at the local level. We place a heavy emphasis on all evaluation activities as providing the local district with some kind of information upon which they can make a decision, look for a direction, or improve something.

The Wyoming Comprehensive Systematic Planning Requirement, which becomes mandatory this coming September, is simply a minimum standard under the Wyoming school code. It is not imposed by the legislature. The requirement states that local districts must conduct needs assessments, they have to be able to demonstrate community involvement in planning, they have to show that they have evolved procedures for resolving identified learner needs, and they have to evaluate. The requirement allows districts to pursue planning as they see fit--with State department assistance, employment of outside consultants, doing it on their own--but to have their plan approved by the State department.

In our needs assessment, as it is currently structured, we start with broad, basic community involvement, having the districts that we work with form a needs assessment committee and a steering committee to monitor the procedures. We have been using as the basis of the needs assessment the Phi Delta Kappa goal statements for education, but in addition to these we ask for identification of learner needs, and let other needs evolve from those. We don't require testing as an initial step in needs assessment. We encourage community input with regard to the goals viewed as being important for education in the local district, and then an estimation of how the district is perceived as currently meeting those goals. Wherever there is a gap between high value of goal and low rating of fulfillment in the district, that becomes a need; and at that point, the district is asked to document the need in more detail, not only by using information from community surveys but also by using any school data that is available. The State Board gives each district discretion to modify the systematic comprehensive planning model. The only requirement imposed is that planning be systematic and comprehensive.

In the planning model the first step is the completion of a needs assessment. The next step we suggest is that the school district take the results of the needs assessment and further define its problem. Analyze the needs assessment and generate some alternative solutions to meet the needs identified. Devise an evaluation scheme for assessing the extent to which the district has met the goals they have set for themselves.

We have been working with the districts, offering technical assistance. Where a given identified need requires some kind of expertise we don't have, we have been drawing on the subject-matter consultants and the instructional division for help, but we haven't been providing special money.

Another facet of the entire planning requirement is tied to the accreditation process of local school districts. We have started accreditation visits which not only involve looking at those items which come as a part of our regular accreditation program, teacher certification, and other related concerns; but as a part of that overall look at the district, we have been re-examining their progress in meeting the planning requirement. We do get a little internal criticism about that and have some reservations about it, but most districts seem to be cooperative.

Another facet of what we are involved in is, of course, evaluation of federally funded programs. I was rather encouraged to hear our keynote speaker suggest backing off from a rigid position of objectivity. Our efforts in Wyoming are certainly not solidly objective. We can't point to any evaluation activities we are involved in that center around sophisticated research designs. At the same time, I think we have been rather direct and honest in using program objectives developed by the local district as a basis for evaluating programs, and we generate our questionnaire and student interview guides and staff interview instruments from objectives for a given program and go in and collect data and direct our generalizations to positive areas and areas of concerns. Recommendations we make are based on what a given district's project said it was going to do in its project objectives. I suppose some people would call our efforts goal-based evaluation, and at times we question if maybe even this isn't being a little dishonest because frequently as we look at a project we see side issues that don't come out of the objectives stated by the project. We are encouraged by the fact that with some of the programs like the Title III programs, some of our recommendations have been taken to heart and some effort has been made to steer the project in a new direction. So we feel as if we are having some effect and doing more than just providing a written report to the federal program director responsible for the project. Of course, I think that is what the whole thing is all about.

As a part of our evaluation activities, we have been working with districts in the area of staff evaluation. We haven't been writing their staff evaluation programs for them, but we have been presenting districts with a model which focuses on self-evaluation techniques and job targets; and we have been working with them to show them the advantages and disadvantages of evaluating staff in different ways and tying this to a responsibility for not only evaluating staff, but also building in some kind of a support system to help them improve.

One of the big concerns expressed by districts, especially in those first few orientation sessions when we start talking about the planning model and we get down to that evaluation component, is who is evaluating and what? What are you going to do to us if we fail? We have been trying to reassure them that we are not passing or failing them on what their evaluation data shows. We want them to become familiar with the process and use it and use the results of their evaluation to make their own decisions, but many of them are certain we are going to penalize them in some way.

EVALUATION OF CAREER EDUCATION IN ARIZONA

by

Richard Ruff

Let me review with you some considerations in developing career education evaluation capabilities at the State department level in Arizona. We analyze management data, student outcome data, and data concerning instructional strategies. We have it done by external evaluators. We try to determine whether the kids are really learning anything or not. We also try to establish the effectiveness of the different approaches that we are using in career education. We have done that in a number of our projects for a number of years, but we do not do it in every project. We simply can't afford it because it is fairly expensive. For example, in Pema County, which is our largest project funded--$600,000--we are putting about $35,000 into the evaluation.

Another type of evaluation we are doing is the Senior Follow-Up Study. We use a different kind of form, specifically geared to try to tap into things we are concerned about in career education. We give it to all the students rather than just vocational education students. We are trying to determine whether the students have achieved any of the outcomes we are after and whether the strategies we are using are effective.

Our Annual Statistical Report is a report we give to the legislators . It's simply number data. We also give them a synopsis from the report. Our Annual Evaluation Report is a report that each of our projects must submit at the end of the year. They have to state each goal, each objective, each evaluation procedure and include their findings. In most cases, that is evaluation they do as part of their projects; it is not done by an external group.

We have also developed a needs assessment system. It's essentially a service to the projects, which gives them an opportunity to find out what the career education priorities are in the community. We will help them at the State department to develop the forms, we do all the data processing, and we give them back a computerized report. It's a system that is very flexible and very simple, so that most projects can do it on their own once we give them some initial instruction. Since we spent a lot of money developing curriculum in career education, we obviously attempted to assess the effectiveness of the curriculum. We did pilot tests, we did field tests, we then went into what we called implementation testing to find out the effectiveness of those units. Now, whenever teachers want to use a unit, the evaluation report tells them how well the kids liked it, how well the teachers liked it, how well the students did on the performance objectives of that unit, and how much it cost to implement.

We are also doing what we call "listening sessions." Essentially this is an opportunity for people in the State department, mainly the director and myself, to go out and really talk to teachers and administrators and project staff to find out what is going on. They are structured interviews. We ask standard questions and then just shut up and listen. This helps us in trying to plan for the subsequent years.

One of the biggest problems I find in evaluation is that most administrators do not understand evaluation reports. Part of my job is to serve as a translator between the evaluator and the administrator. We spend a long time working with

our external evaluators to make sure that they are responsive to the administrators, responsive to the kind of things that the policy or governing board of the project can understand.

In some of our projects we found out that we weren't hitting the high schools nearly as much as the elementary schools. We were putting a disproportionate amount of our staff time into contacting people in the elementary schools for the simple reason that elementary schools, in general, tend to be more receptive to the concepts of career education. Our project staff was going to those places where people loved them. They were welcomed there, so they went there more. They did what they liked to do. We needed data to give us a snapshot of how our staff was spending its time. For example, if the project had a major thrust in parental involvement, and our staff was not talking to anybody about parental involvement, then it is very unlikely that that is going to get done. So we obtained a simple kind of management data on what our staff is doing.

We have also obtained data on the major thrust of speakers' visits to the classroom. Did they deal with the area of career awareness? Did they deal with the area of self-awareness? Which of those areas did they deal with? We found that in elementary schools 73 percent of the guest speakers were dealing with the concept of career awareness to the exclusion of other things. So, the projects made some changes to try to help guest speakers get involved in other kinds of topics. We put together, for example, a little brochure on "Tips for Guest Speakers." The following year, '74-75, we corrected that imbalance to some degree.

(James Mortensen, Colorado, and Helen Warner, Idaho, made presentations on their respective States. Due to mechanical problems with the tape recorder and insufficient time for duplication of additional speeches, we are unable to provide their presentations.)

INVOLVING ELEMENTARY AND SECONDARY TEACHERS IN DATA UTILIZATION

by

Bill Connett, Montana

Lately many people are critical of policy makers and decision makers to use research data. I tend to react negatively to seat-of-the-pants decision-making, but I don't think we can expect people to use data intelligently until they have had some experience and some training in doing it.

I would like to propose that we place more emphasis on teaching people in the elementary and secondary schools how to use information. We cannot expect someone who hasn't had any real training in the use of data to use data suddenly in a rational manner. I don't think that the evaluators and the people who produce the data can do it all themselves, and I don't think we can place all the burden of changing on the evaluator, although I do believe that part of our job is to try to train people in using information. People frequently oversimplify issues because they want simple solutions. That is what they have been used to; that is what is easy. They need to be taught to use complex information in making rational decisions.

EVALUATION ACTIVITIES OF THE PLANNING, RESEARCH, AND EVALUATION SECTION, OKLAHOMA DEPARTMENT OF EDUCATION

by

Charles Sandmann

In Oklahoma, we have an expanded view of evaluation. We know that, technically, evaluation is a decision based upon a comparison. But, we have learned that assumptions cannot always be made about the process leading to evaluation. Our efforts toward evaluation have been influenced by the impact of federal programs as they operate at the State level and in the public schools.

Our activities can be grouped according to the services performed for the department and those for public schools. For the department, we have produced the needs assessments essential for our State plan and the program reports that describe federal programs in our State. We also have been instrumental in the evaluation of workshops conducted by sections of our federal programs division. Perhaps our most important function, and the one which has made the greatest change in the use of evaluation, is staff development. Several years ago, we introduced management by objectives to the department staff using materials from RIPP Conferences. We participated in Evaluation I from UCLA, CSE, and then conducted the workshop for our staff and for LEA personnel.

Accountability has been the second largest impetus to evaluation in Oklahoma. Four years ago our legislature passed a resolution directing the department to implement and monitor an accountability program. They may have had other things in mind, but we designed a three stage process based upon systems analysis. Our experience with MBO provided the basis for our planning. We conducted in-house workshops and then expected the total SDE staff to provide leadership to educators across the state. We produced materials, a slide tape presentation, and a video tape which was used for a simultaneous, statewide workshop for educators. It will come as no surprise to learn that we met with some resistance to the imposing of MBO from the top down.

Our next thrust grew out of our experiences with accountability. We realized that the use of evaluation requires a change in people. To quote John Goodlad, "School change is people change." We selected a people-change technology, Preparing Organization Development Specialist (PODS), from Northwest Regional Lab. A cadre of leaders was developed, and now we are providing the service at the request of our staff and public school staff. It is too soon to describe the effect of OD in our state, but we hope to devise a method for evaluation--although its developers admit that evaluation has been elusive.

The schools in Oklahoma most receptive to both MBO and OD are the schools with federal programs and specifically innovative programs. We provide the data treatment for those who request our services, and these reports are the only ones that could be labeled "evaluation," even though they belong in a comparison phase--they compare measurements made of the project's objectives with a standard defined by the project. Evaluative statements are the prerogative of project personnel and the on-site evaluation team.

Finally in our classification of evaluation services, we have placed the surveys conducted at the request of department or LEA personnel. These are usually

designed to meet a specific need or provide information about a question posed by the person requesting assistance. The following list of publications includes examples of each activity; most are available for your use.

An Assessment of Educational Needs for Students in Oklahoma, Phase II, 1970.
An Assessment of Educational Needs for Students in Oklahoma, Phase III, Vol. 1, and Vol. 2, 1973.
Oklahoma Statewide Assessment, Needs Assessment, Phase IV, 1974.
Migrant Education Handbook, 1975.
Survey of Migrant Education in Oklahoma, 1976. Set of survey forms.
Survey of Migrant Education, Final Report, available October 1976.
Annual Evaluation Report of Federally-Funded Educational Programs, FY 1975.
Evaluation of Education/Fair, 1975.
Evaluation of Education/Fair, 1976.
Evaluation Report for Emergency School Aid Act Workshop for Oklahoma E.S.A.A. Educators, September 27-28, 1975.
Evaluation Report for Value Clarification Workshop for Educators in Southeastern Oklahoma, November 8-9, 1975.
Evaluation Report, Value Clarification Workshop for Educators, Southwestern Oklahoma, April 8-9, 1976.
State Department of Education Goals for the Middle 1970's.
State Department of Education Futures Study, 1976.
Evaluation Workshop I, Center for the Study of Evaluation, University of California at Los Angeles.
Needs Assessment, 1973.
Handbook for Objective Writing and Program Development, 1974-75.
Evaluation, Book I.
Worksheets for a Workshop Presenting Book I.
Affective Measurement, Three Scales.
Group Process Skills, Northwest Regional Laboratory, Eugene, Oregon.
Arts Communication Training, A Workshop for Training Planning Teams for Arts in Education, 1976.

INFORMATION UTILIZATION IN DECISION-MAKING

by

Dr. Roy Forbes, Director
National Assessment of Educational Programs
Education Committee of States

Many times we try to think of "utilization" strategies to make sure that we take the information we have generated and apply it in such a way as to re-plan or re-cycle our educational programs. Whenever we start thinking about information utilization, the first thing that comes to mind is the problem of communication between those of us who are responsible for generating information and those of us who are responsible for using that information in making decisions.

Systematic decision-making does include several steps. The first, obviously, is to define the problem. Everyone claims that step is rather obvious, but I know many times groups who are in the process of making some type of decision only talk about the problem in very general terms and never get around to saying just exactly what the problem is.

Once the problem is defined, then you have to generate a possible solution or possible solutions to that problem. I personally think that you need to do the latter--that is, come up with a number of alternatives so that you can then mull over each one. Do something like identify financial resources or personnel resources that you would require in order to implement one of those possible solutions or identify any problems you would have with implementing one of them--try to identify the constraints--try to identify possible ways of removing those constraints by working around them. Then look at the information you have at that point, reviewing the possible solutions for potential payoff--which one will have more beneficial effects than the others. You have to continually go back to the previous step because usually what your head is doing is synthesizing, combining those possible solutions that you weighed the whole time; and each time you do that, you should go back and create the information needed for each new possible solution. At some point you have to make a selection--the last step.

These are, very briefly, the steps to follow in trying to reach a decision, in trying to solve a problem. Each step requires certain amounts of information selection. Sometimes we can sit down, when we start, and find all the information that we are going to need ahead of time; but in most cases we need to go back and work through each step, identifying additional information, which we can then use in this systematic process.

Sometimes the best decisions you reach are those decisions you make in the shower or driving to work, over a drink, or dancing with your wife--they just pop into the mind all of a sudden. When you stop and think about that particular decision, oftentimes you can stick with it. I call this model the *Ah hah!* type and make the best use of it I can. Now some people might say that is being very unsystematic. I really don't know because I don't know what's happening in the brain at the time that these *Ah hah's* occur. Maybe a thousand years from now we will know how to analyze thinking better than we do today, and someone will be able to determine a more systematic process for that.

One thing you can do which ties in with utilization of decision making--whenever you have one of those *Ah hah's*, sit down and try to reconstruct what information you believe you have used that led you to that decision. I think that's the clue. Many times we feel that we haven't used information in a very systematic way in reaching a decision that way. Sit down and try to figure out all of the processes that you could have brought into play to reach that decision, and you will start identifying information that you may have subconsciously used. Then you can identify those pieces of information and consciously use them.

I think we need to develop a respect for information which appears in a different way from the systematic way. If you do not go back and look at the information that you used, maybe subconsciously, or look at the information based on experience, there is a danger that you will think you reall[illegible]on't need other information, and it creates an attitude about the use of [illegible]nforma-tion that I think can be very damaging.

That gets us to something I call "justification." If someone has a good feeling about a particular decision, or they have a decision that they hope is going to be made, then they will go looking for data to justify that particular decision. In doing so, they are probably going to bump into some information that is not completely in support of their decision, and they will probably change their decision a little bit. They will bump into a lot of information that says they are all wet. They will have to be changing, improving, proving--but I think their finding the justification for a decision is really going to help them gain courage and not make them feel guilty about using information to justify a decision they have made.

Another way you can make sure that justification turns into a plus and not a crutch is by asking yourself and your staff to list the pro's and con's of the recommendation. If they recommend a particular decision, ask them why it shouldn't be done that way. Initially, you will probably shock people, but they learn rather quickly that you are making sure all the information has been considered. We are back to justification again, but we've added a new dimension to it, in that we are building in the requirement that they look for information that would have a negative impact on the decision as well as those things that support the decision.

I have been talking in general terms. Now I would like to talk more specifically of the data itself. One time I worked on a report which showed the progress being made in a particular school in Louisville as a result of a number of innovations. The report referred to the vandalism rate in the school. During the middle of the year we began to get very excited about our data because we saw we were having a tremendous drop in the vandalism rate, and we thought that was something extremely positive. We started patting ourselves on the back a little bit. In several more months the rate went down even more, and then we were convinced that we had something good going until we talked with the maintenance man. We had determined vandalism by glass breakage until we discovered that the maintenance people were putting in plexiglass that wouldn't break. A drop in glass breakage didn't mean a drop in vandalism rates. It just changed the way vandals operated a little bit. One of the dangers in dealing with data lies in not taking a second look.

Before we get too excited about data or start using data for making decisions, we owe it to ourselves to go back and take a good hard look at the data. That is a very important step in utilization of data. You have to sit down and identify both the strengths and weaknesses of the data to determine the limitations of it--to know just how far you can go with it. For example, National Assessment data cannot be used for cause and effect information. That's a misuse. All of you can think of misuses of data. I think the current misuse of the ACT and SAT scores by some people is a good example. These two tests are designed to depict the success of college freshmen and they do a very good job of predicting the success of freshmen, but people try to take those tests and translate them to see how well we're doing other things in education. It is foolhardy not to consider the weaknesses of data--not weaknesses in terms of what they were designed to do, but weaknesses in terms of what some people are trying to use the data for.

Another mistake we make with data is to dismiss it rather rapidly. It would be very easy for someone to make the same comment about achievement data for the school system as I just made about the SAT and ACT data: "Let's not use it because it has a flaw in it; it is intended for another purpose." That is just as serious an error in data utilization because we have to look at data to determine what use we may have for it. So, if someone from your staff rejects a piece of data that you think they should use in reaching a decision, then again ask why it should be used. Very little data generated has absolutely no value.

It is also important to find out how we feel about the data. Do we like it because it supports our position, and that is reason enough for using it? Do we tend to disregard it because it tends to be a threat to us, suggesting some things that we would just as soon not hear about? Are we a little bit defensive about the data? We have to ask ourselves all of those questions to determine if we have some barriers in making use of data that could be extremely helpful--barriers of insecurity that sometimes occur when we look at data that doesn't please us. It is extremely important to identify the strengths and weaknesses of the data that you project to use and to have an open discussion about the way you and your staff feel about that data.

When you go through that kind of process, you can discover the good things and the bad things about the data. You can then communicate that information back to whoever produced the data--because the producers of data, like National Assessment, State or government programs, assessment or evaluation programs, are all dependent on the utilization of the data in the decision-making positions. Everyone in the job of producing data should welcome constructive criticism; but if you go through a process of identifying strengths and weaknesses of data and keep these to yourself, you are not helping yourself. As far as National Assessment is concerned, if you let us know ways in which you consider the data to be weak, then we can respond to those weaknesses; and if you will let us know ways in which you consider the data to be strong, then we can be sure that we don't do anything to threaten the strength of the data. But you have to communicate to data producers a need to change what they are doing, improve what they are doing, or forget what they are doing and stop doing it.

Any time we make use of data we make certain assumptions about what the data reveals. It is a wise idea to check out your assumptions with the producers

of the data so that you feed back information and check out your assumptions. Maybe the producer did not realize that people were making the assumptions you did or that perhaps there was something else that should have been considered.

As you start developing a personal contact with the generator of the data you use in decision-making, it is extremely important that you become an owner of that data. There are a couple of ways to do that. I have already mentioned some of them. Identifying the strengths and weaknesses of the data requires an inquiry into the data which causes you to start feeling more comfortable with it and it starts becoming a part of you; you now own it because it is something that you have working for you. Listing the pro's and con's of some position that you are taking is another way to develop ownership because the data are in some way supportive of where you are coming from. Identifying some ways in which even what you consider to be weak data can be used, even that process develops ownership. Personal contact between the generator of the data and you or your staff adds another dimension to that ownership. Probably the most critical thing is that if you are going to use the message that the data generator used in generating data for yourself, you ought to not replicate it--you ought to refine and improve upon either the methods of collecting the data or try to improve the data base by adding additional information from other sources. That probably goes further to create ownership in the data than anything else.

Finally, we have to watch out not to over-simplify the use of any data. Many times people expect of us to have either yes-no or either-or types of responses. They don't want to deal with complexities, and so it is very easy for us to look at the data, over-simplify, and try to come up with those yes-no's, up-down's, either-or's. We're all probably guilty of this mistake at some time; so as a final step in the utilization of information in decision-making, we have to inquire whether we have tried to over-simplify or whether we have dealt with the complexity of the data in order to make the decisions that will help us do a better job in education.

Then, when we are faced with momentous decision-making, we can reflect on the decisions that we have already made and the information that was used in making those decisions, and we will realize that, yes, indeed, we do make use of information in trying to replan new educational programs. And, by using some of the strategies that I have mentioned this morning, hopefully we will create a desire in the people who will have to make use of the information to go out and search for good information and then to make the best use of it they can.

SUMMARY, EVALUATION AND IMPLEMENTATION

by

Gerald Kowitz, University of Oklahoma

Views of the Evaluative Process

One of the central problems facing this group is that nobody likes to be evaluated--nobody. There are some of us who take great glee in evaluating others, but when it is our turn to get under the microscope, it hurts.

We have several problems. That means we must begin with some axioms, things that we accept. If we are going to get something done, we have to begin somewhere with what we believe. And for our axioms, I propose the following: accountability is inevitable. If you're spending somebody else's money, they are going to call you to account. Accountability presumes assessment plus judgment. This should lead to improvement. Improvement assumes control. This is a very important axiom because it has as a corollary, a non-symmetric: control is based upon knowledge. Control based upon knowledge has the non-reversible, non-symmetrical corollary that controls cannot give us knowledge. This is extremely important for evaluation. If we can report no difference, what we're really saying is that we didn't find any difference. We can't say there was no difference.

We've only begun talking about policy. The history of studies of policy are not yet a decade old. Oh, we'll find philosophy on them. You can probably go back to the first great policy-maker in history, Joseph, who saved Egypt through his policy. Or if you prefer, the story of Jethro who'd been to group dynamics when he told Moses, "You can't be bothered with all these problems. Break them up in groups of ten, tell them to go out and discuss it themselves, and then come back and report." You don't believe me? Look it up in Exodus. That's what he told them. That's how far we've come. We talk about policy as a means of control. We have control in physical sciences. We are marveling at our landing on Mars. We got it there. It worked. When you think of the marvelous things that science can do I wonder why it is that they can't build a smooth railroad crossing. Why is it that building ventilation is never quite right? Is this so much more difficult than putting a man on the moon? Or a spider on Mars? Nobody answers. We have the idea of controlling people and society. We have philosophy, traditions, law, the social sciences. We usually think of them as less precise than the physical sciences. I'm not so sure.

Finally, we are now getting into the study of the control of the control. This is the study of policy. We are not talking about ethics or philosophy--philosophy and ethics are respectable fields in their own right. We are talking about control of the controller of policy, the policy maker. At the present time, we see this primarily as a social, political manipulation which, among us, is flipped off in the phrase RFP, request for proposal. Somebody has made a policy and they want us to do it. I've often wondered, "Why don't you do it yourself?" When we do get into this arena of social, political manipulation, data becomes ammunition for guerilla warfare. Evaluation plans become the strategies and tactics of trench warfare. I would suggest to you that one of today's great differences between an educational researcher and an evaluator is that the researcher is looking for solutions to problems. The evaluator is

trying to arrange solutions that will work, using the problems that have been solved, the solutions to the problems.

We talk about evaluation with two faces. We're looking in two directions and we're seeing two different things. One, we are trying to justify the past. This is good; it has advantages. We should substantiate what we have done. Call it justification--that's all right. But let us remember that we can't change the past. The second role of evaluation is predicting for the future. This will help us with our plans.

Objective-External Evaluation

There are roughly three different views of the evaluative process. The first I like to call objective-external evaluation--the optimum, the thing that we all wanted to do for years and years. Yet many of us, who have thought about it in the dark of the night when we couldn't sleep, realized we can't get there. This is scientific research at its best. Objectivity, external, no connections, and evaluative. It is predictive. If you do this, then this will happen. One of the bases of this evaluative process is behavioral objectives. You can not do objective-external evaluation without behavioral objectives. But just because you have behavioral objectives does not mean that you have objective-external evaluation.

For example, last spring I was asked to work at the Local Association of Sunday-School Superintendents in developing and evaluating their new curriculum projects. By the time I arrived, all the Sunday school teachers were writing behavioral objectives. I was to evaluate them. They put me up on the podium and started handing them to me one at a time. I was doing fairly well with the idea that after so many sessions the children would know and could name at least the 12 disciples of Jesus. Then I came down to one that this little old lady had written. It said, "Of the children that attend my Sunday school class 80 percent of the time, 50 percent will go to heaven."

Integral to this idea, remember, is it's external. This means the evaluator desires as little involvement in the program as possible. He does not want to be confused by the facts. He knows what he is doing. It has the great advantage of permitting you to be creative. It is supremely defensible, it is mathematically justifiable, it is totally logical, and it has extremely high generalizability because there is great focus on the problem at hand. People working in this area are not interested in 100,000 cases, because they know that any test statistics that they apply will be of some significance if you have enough cases. If you get enough cases, you will get a statistically significant difference even though it doesn't mean anything.

Sampling cannot be well controlled. We all know that. Politics frequently interfere with the design. There is a great temptation to generate case studies, resulting in almost total loss of generalization. The evaluator knows what is right, and would do right, but cannot do right because of the political, social implications.

Subjective-External Evaluation

Next, I would like to discuss subjective-external evaluation. Here we are getting the judgment of experts who deliberate on how things ought to be. Now,

here is the difference. When we were talking about the external-objective, there is an assumption that policy will automatically flow from experimental results. You can't deny it. If we find these facts to be self-evident, that's it. With subjective-external evaluation, we are going down to allow a little elbow room.

Experts will set up standards. They will dictate to you some policy. They will expect you to live by it. The operators in the program do not set the quality control objectives--they are set by the experts. The program directors have to live with them somehow. This is a problem that is locally known as the fallacy of Piaget. Before you can conduct any Piaget experiment, you have to be trained by certified Piaget trainers. They give you your glasses, and looking through those glasses, you can only see what they have seen. Here you have a very real problem. You are merely saying, yes, you saw what you saw. But that might not be right.

The evaluator's role in subjective-external evaluation is that of the superior expert. We have put together the best brains we can find, and we've given control to the top level. If we didn't select the right people, we're in trouble. This is a very dangerous, although useful, weapon. The evaluators or experts can come in and criticize you for things you never tried to do. They will totally overlook things that came out of your plan that don't fit their preconceptions. In fact, you will find that this is one of the best methods if you want to play hatchet person. The problem is it lacks generalizability.

Internal-Developmental Evaluation

Finally, we had a newcomer to the arena, only about 40 years old now, internal-developmental evaluation. This is only one component of the program. The evaluator's goal here must be perfectly congruent with that of the manager, that is, to improve the program. We are not talking about achieving specific objectives--what we are talking about is not what the organization can do now or what it has done in the past, but what is its potential for future performance and how can we expand and increase that potential. This requires that the evaluators must be involved from conception to termination.

Systems analysis is a great tool with this method. We're talking about an evolutionary operation. We can't shut down the school to improve it. We have to introduce small bits and see how they work. We will take out throwaways and put others back in. But we are going to be working together as a team, evaluator and program director, in this repeating process. The advantages here are the involvement and the motivation for the people concerned. If not motivation at least consider you are a part of an important group doing things.

The disadvantages are very interesting. The greatest disadvantage is that you are dealing with highly abstract ideas. You may never get down to touch earth. Your behaviorally-stated objectives are of little help. You will be restricted by the policy guidelines of the organization. You must be consciously political, and you had better be good at it. You may become, unfortunately, extremely gimmicky--such as getting people together to role-play, wearing masks and costumes, or playing at various kinds of games--in-basket, out-basket. We have here the Delphi Technique that more frequently than not bypasses the best solutions because nobody wants to talk about them at first; so they are voted

down right away. We have assertiveness training, which frequently deteriorates into attack training. Evaluation is definitely post-hoc, and lacks generalizability.

Reactions to RIPP Program

In my final reaction to the RIPP program, I would stress the importance of meeting together face to face as people. Words said to each other have much more meaning than the vibrations that come over the telephone. It is much easier to say no over the telephone than it is face to face. It is much easier to lose a letter in the in-basket or the waste basket than it is to give it a possible answer. It is very important to see people, to know people.

If you are going to talk about your organization as a region, you have got to get to know that region. This means somewhere in your conference I would strongly recommend you schedule a regional familiarization activity. People are going to do it anyway. If you can do it with new friends, particularly from the area, you will begin to get a regional identification and consolidation that will do your organization more good than a very large federal grant without assessment attached. Site visits are extremely important.

I would suggest that we need to watch our language when we start talking about how evaluation is too complex for those poor dopes out there. Don't assume that they are uneducated. They are unwilling to be educated. Maybe they need a little nourishment before they will throw themselves into it. I suggested to you three models. Please do not consider them as alternatives. And don't think maybe we can string them together as one--each has its purpose, its advantages, its disadvantages.

EVALUATION FOR EFFECTING EDUCATIONAL FUTURES

Seminar Evaluation Form

I. DEMOGRAPHIC:

1. Professional Position (Check One):

12 local school district personnel

31 state education agency staff

3 education service center personnel

1 higher education faculty or staff

1 school board member

1 other: Contractor

2. What State do you represent?

II. OBJECTIVES:

3. Were you able to identify probable directions of change in education and ways in which evaluation could assist with the change process?

positively	yes, generally	only somewhat	not at all
4	28	14	3

4. Were you able to share experiences and ideas on evaluation activities and findings?

positively	yes, generally	only somewhat	not at all
19	24	6	-0-

5. Were you provided the opportunity to review promising or emerging evaluation models and practices?

positively	yes, generally	only somewhat	not at all
23	19	7	-0-

6. Did you gain ideas and suggestions that could be useful in your situation or responsibility?

positively	yes, generally	only somewhat	not at all
18	20	11	-0-

III. MAJOR ACTIVITIES: (rate the following activities; circle your response per activity in each of the three areas noted below)

	KNOWLEDGE I GAINED				POTENTIAL APPLICATION					PRESENTATION TIME			
	a lot	a fair amount	not much	no comment	will use	may use	won't use	not applicable	no comment	more needed	OK	too much	no comment
7. RMC Research Models	16	28	3	2	13	22	5	5	4	7	34	3	5
8. Abt Assoc. Inc. Models	4	12	29	4	2	13	20	8	6	-0-	33	11	5
9. Calif. Eval. Project.	17	25	4	3	10	27	6	2	4	3	38	3	5
10. States Show & Tell	25	22	2	-0-	20	23	2	-0-	4	5	35	5	4
11. Utilization of Data to make decis. ECS.	11	15	22	1	14	15	7	6	7	3	40	4	2
12. Systems Development Corporation	11	23	3	12	6	24	4	2	13	5	30	2	12

13. Note any comments you may have concerning the presentation:

a. RMC

Quiet Technical	2
Should have been questioned further	2
Not well presented	3
Well prepared & professionally presented	30
No Comment	12

b. Abt

Not well presented	6
Poorly planned project of research	3
Well prepared and presented	5
Waste of time	4
Too much rambling	3
No Comment	28

c. Calif. Eval. Project

Well prepared and presented	5
Missing in tone	3
Good practical approach	26
Good/Fair	4
No Comment	11

d. States Show and Tell

Well prepared and presented	8
Too many presentations atone time	3
Excellent opportunity to compare and share	25
Very Poor/Fair	2
Good	3
No Comment	8

e. Utilization of Data

Did not get into depth of problem	3
Well prepared and presented	4
Fairly informative	28
Worthwhile	3
Fair	2
No Comment	9

f. Other Events

Interesting Conference - Well Organized	22
Wrap-up Evaluation - O.K.	20
Speaker was Superb	2
U.S.O.E. Presentation - valuable and useful	5
No Comment	2

IV. OVERALL RATING:

14. In general, how applicable were seminar activities to your local situation? (circle one)

highly applicable	generally applicable	not applicable	no comment
11	33	1	4

15. Conference organization and implementation was . . .

superb	good	so so	terrible	no comment
15	30	1	-0-	3

16. The conference menu of activities and events was . . .

steak & potatoes	steak	potatoes	hors d'oeuvres	beans	other	no comment
13	20	5	3	-0-	O.K. Jim Bean Hamburger & Potatoes	8

V. FOR THE FUTURE:

17. What should the conference planners have included on the agenda that they did not include?

More direction for policy makers	3
More "local school district" representation	4
More "local school district" presentations	3
More panel discussions	2
More emphasis on what actually is going on at state and local levels	6

Good job	2
Good effort	3
U.S.O.E. reactions	19
More interaction with groups	3
More time on the Calif. Evaluation	
Training package	4
No Comment	18

17. Anything Else?

Field Test Future Forms	2
Group exercises well liked	3
Panel most informative and useful	5
More real discussions	1
Good conference	6
To intense - more varied	2
To much time	1
Final summary overview - excellent	29
No Comment	0

VI. COMMENTS ON KEYNOTE SPEAKER:

1. Excellent, set right tone for the entire conference.
2. Very interesting.
3. "Heavenly minded" - only one point retained-importance of process evaluation couched in lofty terms.
4. Very good and controversial.
5. Very interesting, agreed with my biases/convictions so, therefore, did not introduce many new ideas but did serve to support and facilitate my own thinking on many issues.
6. Enjoyed the speaker very much. Did not learn much that I didn't already know but he did serve to refresh some philosophical ideas that set the scene for the conference.
7. Somewhat removed from the real world of where the interfact between kids and programs occurs.
8. He was a good speaker. Tried to get across a balance between objective and subjective assessment. Perhaps a little too general.
9. Very perceptive about future needs regarding education.
10. Afraid I didn't get much out of it - lots of platitudes etc.
11. Excellent choice and did an excellent job. He was "down to earth".
12. Stimulating, controversial, got us thinking and new ideas.
13. A small, small peek through a tiny crack in the evaluation door. Humor, A+.
14. Dynamic; set the right tone. Easy to follow and humorous.
15. Average.
16. Very good, brought good information.
17. I appreciated the keynote address primarily because of its "scholarly" tone, which probably isn't appropriate for a summary. This is the first session, conference, meeting, etc., I have attended which had a single scholarly address; it was great. I was beginning to think (after 7 months with the State Department of Education) that intellectualism was dead among "professional educators".

VII. SUMMARY SPEAKER:

1. "Earthly Good". Entertaining, informative, and excellent debriefing which wove the various experiences of the session as threads in a tapestry.
2. After hearing the Summary Speaker, I had a much better idea of what the Keynote Speaker said - I understood the summary version best.
3. Very comprehensive and perceptive coverage of the topics/issues raised during the conference; very stimulating presentation; good speaker; share important insights into the evaluation process.
4. I enjoyed his practical approach to evaluation. His humor added to the presentation. I agreed with his ideas that some speakers were patronizing and that the program should have had more time for sharing projects and ideas - excellent.
5. Closer to the real world of the classroom.
6. Entertaining but took many easy cracks that were unjustified. Didn't really have anything to say.
7. Very good summary of conference.
8. One of the finest wrap-ups I have ever heard. Glad we can laugh at ourselves and share some serious thoughts too. It's nice to have someone give a pretty objective commentary on the state of the art.
9. Very entertaining and "on target". He did an excellent job on putting the speakers in perspective.
10. Went beyond summary to contribute much itself. Good, on target - could have been just a little briefer.
11. Insightful. Suggestions for meeting formats were those I can support.
12. More exciting and well planned summary than I have ever heard. It was exciting, humorous, well planned and to the point.
13. Very good - excellent.
14. Enjoyable and informative - very good.
15. Superb - excellent.

VIII. BOTH SPEAKERS:

1. Both the Keynote and Summary Speakers were excellent. I particularly liked the Summary Speaker as he injected many cute stories and information.
2. Both are men of impressive intelligence and provocotive imagination - excellent.
3. Best parts of conference. Demonstrated more perspective on evaluation than others combined. Mature, professional educators - provided enthusiasm and hope for work.
4. Both Speakers were very good and controversial.
5. Both Speakers offered an appropriate spark - the initial spark to excite; the concluding spark to sustain interest. Please print remarks. Both had great points that merit further thought.
6. Both were interesting, very informative - excellent.